OM nucleic - nucleic search, using sw model

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2005 Compugen Ltd.

Run on:     May 15, 2005, 10:35:37 ; Search time 13.1953 Seconds
                                     (without alignments)
                             9421.101 Million cell updates/sec

Title:          US-10-764-075-2
Perfect score:  21
Sequence:       1 atgaccacacaaggcagatgg 21
Scoring table:  IDENTITY_NUC
                Gapop 10_0 , Gapext 1.0
Searched:       4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters:     8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      N_Geneseq_16Dec04:*
        1:  geneseqn1980s:*
        2:  geneseqn1990s:*
        3:  geneseqn2000s:*
        4:  geneseqn2001as:*
        5:  geneseqn2001bs:*
        6:  geneseqn2002as:*
        7:  geneseqn2002bs:*
        8:  geneseqn2003as:*
        9:  geneseqn2003bs:*
       10:  geneseqn2003cs:*
       11:  geneseqn2003ds:*
       12:  geneseqn2004as:*
       13:  geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.		Score	Query Match %	Length	DB	ID	Description
1		21	100.0	29751	12	ADJ39000	Adj39000 SARS coro
2		17.8	84.8	1560	6	ABV78222	Human MMP
3		17.8	84.8	1560	6	ABZ35798	Human MMP
4		17.8	84.8	1560	6	ABX10041	Human MMP
5		17.8	84.8	1560	6	ABL91763	Human pol
6		17.8	84.8	1818	6	ABV78191	Human MT4
7		17.8	84.8	1818	6	ABZ35767	Human MT4
8		17.8	84.8	1818	6	ABX10010	Human MT4
9		17.8	84.8	1818	6	ABL91732	Human pol
10		17.8	84.8	2295	3	AAA13389	Human MT4
11		17.8	84.8	2295	3	AAA14351	Human MT4
12		17.8	84.8	2306	4	AAH28234	Nucleotid
13		17.8	84.8	2306	10	ADE84832	Farnesyl
14		17.8	84.8	2423	3	AAA14340	cDNA enco
15		17.8	84.8	2423	8	ABZ20920	MT4-MMP C
16		17.8	84.8	2438	12	AAA13377	Human MMP
17		17.8	84.8	2438	10	ACF63409	Human MMP
18		17.8	84.8	2438	12	ADK14165	Human aut
19		17.8	84.8	2438	12	ADO28632	Human MMP
20		17.8	84.8	2438	12	ADQ19326	Human sof
21		17.8	84.8	3007	12	ADQ23578	Human sof
22		17.4	82.9	3003	11	ADM01725	Human cDN
23		17.4	82.9	6952	10	ADL24806	Intestina
24	c	16.8	80.0	74	10	ACD95453	Human col
25	c	16.8	80.0	74	10	ACD93726	Human col
26	c	16.8	80.0	224	8	ABX88992	Corn ear-
27		16.8	80.0	270	6	ABV96477	Human pan
28	c	16.8	80.0	415	2	AAX30910	Streptoco
29		16.8	80.0	451	10	ADD71720	Human uri
30		16.8	80.0	487	2	AAZ96487	S. pneumo
31		16.8	80.0	630	13	ADQ57025	Novel can
32		16.8	80.0	890	10	ABX07482	S. pneumo
33		16.8	80.0	891	12	ADM91971	S. pneumon
34		16.8	80.0	894	4	AAS55733	Streptoco
35		16.8	80.0	894	8	ACA50075	Prokaryot
36	c	16.8	80.0	978	10	ADD71721	Human uri
37		16.8	80.0	1222	3	AAC59283	Human sec
38		16.8	80.0	1323	3	AAC81772	Human bea
39		16.8	80.0	1446	12	ADO07809	Human pol
40	c	16.8	80.0	1792	10	ADD71722	Human uri
41		16.8	80.0	1792	12	ADI26142	Human cDN
42		16.8	80.0	1814	13	ADP24393	PRO polyp
43		16.8	80.0	2354	6	AAD32038	Human kin
44		16.8	80.0	2446	2	AAX89852	Human pro
45		16.8	80.0	2488	4	AAF44675	Novel pro

ALIGNMENTS

```
RESULT 1
ADJ39000
ID   ADJ39000 standard; DNA; 29751 BP.
XX
AC   ADJ39000;
XX
XX
DT   06-MAY-2004  (first entry)
DE   SARS coronavirus nucleotide sequence.
XX
KW   small interfering RNA; siRNA; modified ribonucleotide;
KW   viral replication inhibition; hepatitis C virus; HCV; hepatitis C;
KW   antiinflammatory; hepatotropic; virucide; hepatitis A virus;
KW   hepatitis D virus; hepatitis E virus; Ebola virus; influenza virus;
KW   rotavirus; reovirus; retrovirus; poliovirus; human papilloma virus;
KW   metapneumoniavirus; coronavirus; viral infection; gene; ds.
OS   SARS coronavirus.
XX
XX
PN   WO2004011647-A1.
XX
PD   05-FEB-2004.
XX
PF   25-JUL-2003; 2003WO-US023104.
XX
PR   26-JUL-2002; 2002US-0398605P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Han J, Seo MY, Houghton M;
XX
DR   WPI; 2004-143862/14.
XX
PT   New RNase resistant small interfering RNA, useful for treating viral
PT   infections, e.g., hepatitis C, influenza virus or coronavirus infection.
XX
PS   Example 10; Fig 3; 74pp; English.
XX
CC   The present invention describes a small interfering RNA (siRNA) which
CC   comprises a modified ribonucleotide, where the siRNA is resistant to
CC   RNase and retains the ability to inhibit viral replication. Also
CC   described: (1) inactivating a virus in a patient; (2) making a modified
CC   siRNA that targets a nucleic acid sequence in a virus; (3) a double-
```

CC stranded RNA molecule of 10-30 nucleotides that inhibits replication of
CC hepatitis C virus (HCV); (4) inducing targeted RNA interference toward
CC HCV in hepatic cells; (5) inhibiting replication of HCV; (6) a vector
CC comprising a DNA segment encoding the RNA molecule; (7) a host cell
CC comprising the vector of (6); (8) inhibiting replication of HCV in cells
CC carrying HCV; (9) treating hepatitis C in a subject; (10) a modified
CC siRNA molecule comprising a double-stranded RNA molecule of 10-30
CC nucleotides in length, which mediates RNA interference toward a target
CC agent or virus and is linked to at least one receptor-binding ligand; and
CC (11) inducing targeted RNA interference in a patient. The modified siRNA
CC molecules have antiinflammatory, hepatotropic and virucide activities.
CC The modified RNA molecules are useful for inactivating virus in mammalian
CC cells. The siRNAs are useful for treating hepatitis C virus, hepatitis A
CC virus, hepatitis D virus, hepatitis E virus, Ebola virus, influenza
CC virus, rotavirus, reovirus, retrovirus, poliovirus, human papilloma
CC virus, metapneumoniavirus or coronavirus infections. The methods of the
CC invention can be used to correct or compensate for cellular physiological
CC abnormalities involved in conferring susceptibility to viral infections
CC in patients and/or alleviate symptoms of a viral infection in patients.
CC The present sequence represents the SARS coronavirus nucleotide sequence,
CC which is used in an example from the present invention.
XX
SQ Sequence 29751 BP; 8478 A; 5941 C; 6188 G; 9144 T; 0 U; 0 Other;

Query Match   100.0%; Score 21; DB 12; Length 29751;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGACCACCAAGGCAGATGG 21
          |||||||||||||||||||||
Db 29398 ATGACCACCAAGGCAGATGG 29418

RESULT 2
ABV78222
ID ABV78222 standard; DNA; 1560 BP.
XX
AC ABV78222;
XX
DT 15-NOV-2002 (first entry)
XX
DE Human MMP17 DNA SEQ ID NO 106.
XX
KW RNA inhibition; dsRNAi; gene expression inhibitor; oncogene; cytostatic;
KW virucide; protozoacide; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200255693-A2.
XX
PD 18-JUL-2002.
XX
PF 09-JAN-2002; 2002WO-EP000152.
XX
PR 09-JAN-2001; 2001DE-01000586.
PR 26-OCT-2001; 2001DE-01055280.
PR 29-NOV-2001; 2001DE-01058411.
PR 07-DEC-2001; 2001DE-01060151.
XX
PA (RIBO-) RIBOPHARMA AG.
XX
PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
DR WPI; 2002-590671/63.
XX
PT Inhibiting expression of target gene, useful e.g. for inhibiting
PT oncogenes, by administering double-stranded RNA complementary to the
PT target and having an overhang.
XX
PS Claim 10; Page 176; 203pp; German.
XX
CC The invention relates to inhibiting expression of a target gene (I) in a
CC cell by introducing an inhibitory RNA (dsRNA) having a double-stranded

CC structure of at most 49 consecutive bases. At least part of one strand
CC (as1) of dsRNAi is complementary to (I) and at least one end of dsRNAi
CC has an overhang of 1-4 nucleotides. The method is used to inhibit the
CC expression of a wide range of genes, e.g. oncogenes, cytokine genes etc.
CC in humans, also genes in Plasmodium or in viruses or viroids that are
CC pathogenic for humans, animals or plants. Introducing an overhang into
CC dsRNA greatly increases effectiveness for inhibiting gene expression,
CC both in vivo and in vitro and also increases stability and thus the
CC effective concentration inside the cell. The present sequence is that of
CC a gene related to the invention
XX
SQ Sequence 1560 BP; 275 A; 549 C; 479 G; 257 T; 0 U; 0 Other;

Query Match   84.8%; Score 17.8; DB 6; Length 1560;
Best Local Similarity 90.5%; Pred. No. 87;
Matches 19; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 ATGACCACCAAGGCAGATGG 21
         |||||||||| ||||| ||||
Db 1112 ATGACCACGAGGCACATGG 1132

RESULT 3
AB235798
ID AB235798 standard; DNA; 1560 BP.
XX
AC AB235798;
XX
DT 07-FEB-2003 (first entry)
XX
DE Human MMP17 polynucleotide SEQ ID NO 106.
XX
KW Double stranded RNA; dsRNA; RNAi; RNA inhibition; cytostatic; virucide;
KW protozoacide; gene expression; antisense; tumour; infection; Plasmodium;
KW virus; viroid; anti-GFP; human; HIV; human immunodeficiency virus;
KW Hepatitis C virus; human papilloma virus; gene; ds.
XX
OS Homo sapiens.
XX
PN DE10100588-A1.
XX
PD 18-JUL-2002.
XX
PF 09-JAN-2001; 2001DE-01000588.
PR 09-JAN-2001; 2001DE-01000588.
XX
PA (RIBO-) RIBOPHARMA AG.
XX
PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
DR WPI; 2002-683450/74.
XX
PT Inhibiting expression of target genes, useful e.g. for treating tumors,
PT by introducing into cells two double-stranded RNAs that are complementary
PT to the target.
XX
PS Claim 13; Page 73-74; 100pp; German.
XX
CC The invention relates to inhibiting expression of a target gene in a cell
CC by introducing at least two oligoribonucleotides (dsRNAI and II), both
CC with a double-stranded (ds) structure of at most 49 sequential nucleotide
CC pairs. At least part of one strand (S1, S2) of the ds structures in each
CC of dsRNAI and II are complementary to regions in the target gene. The
CC method uses antisense inhibition of gene expression using double stranded
CC RNA inhibition (RNAi). The method is particularly used to treat tumours
CC or infections, especially by Plasmodium or viruses/viroids (pathogenic on
CC humans, animals or plants). The method provides more effective inhibition
CC of expression than known methods using a single dsRNA, even at very low
CC concentrations. When dsRNA has at least one unpaired nucleotide at the
CC end, stability (and thus effective concentration in the cell) is improved
CC and efficiency can be increased further by pretreating the cells with
CC interferon. The present sequence is that of a target DNA of the invention

XX
SQ   Sequence 1560 BP; 275 A; 549 C; 479 G; 257 T; 0 U; 0 Other;

Query Match   84.8%;   Score 17.8; DB 6;   Length 1560;
Best Local Similarity   90.5%;   Pred. No. 87;
Matches   19; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 ATGACCACCAAGGCAGATGG 21
            ||||||||| ||||| |||
Db   1112 ATGACCACCACGAGGCACATGG 1132

RESULT 4
ABX10041
ID   ABX10041 standard; DNA; 1560 BP.
XX
AC   ABX10041;
XX
DT   23-JAN-2003 (first entry)
XX
DE   Human MMP17 DNA fragment SEQ ID 106.
XX
KW   Oligoribonucleotide; interferon; oncogene; cytokine; Id; developmental;
KW   prion; inhibition; human; ds.
XX
OS   Homo sapiens.
XX
PN   DE10100587-C1.
XX
PD   21-NOV-2002.
XX
PF   09-JAN-2001; 2001DE-01000587.
XX
PR   09-JAN-2001; 2001DE-01000587.
XX
PA   (RIBO-) RIBOPHARMA AG.
XX
PI   Kreutzer R,   Limmer S,   Rost S,   Hadwiger P;
XX
DR   WPI; 2002-742209/81.
XX
PT   Inhibiting expression of target genes, e.g. oncogenes, in cells, by
PT   introduction of complementary double-stranded oligoribonucleotide, after
PT   treating the cell with interferon.
XX
PS   Disclosure; Page 78-79; 98pp; German.
XX
CC   This invention describes a novel method for inhibiting expression of a
CC   target gene by introducing into the cell that contains the target gene at
CC   least one oligoribonucleotide (dsRNAi) that has a double-stranded (ds)
CC   structure of not more than 49 consecutive nucleotides (nt), where at
CC   least a segment of one strand of the ds structure is complementary with
CC   the target gene and the cells are treated with interferon before
CC   introduction of dsRNAi. The method is used to inhibit expression of
CC   target genes, particularly oncogenes, cytokine genes, Id (not defined)
CC   protein genes; developmental or prion genes, or genes expressed in
CC   pathogenic organisms (particularly plasmodia) or in viruses or viroids
CC   (pathogenic in humans, animals or plants). Treating the cells with
CC   interferon greatly increases the extent to which dsRNA can inhibit
CC   expression of the target genes, and the effect is even greater when dsRNA
CC   are modified to increase their stability. ABX09936-ABX10075 represent
CC   gene fragments used to illustrate the method of the invention
XX
SQ   Sequence 1560 BP; 275 A; 549 C; 479 G; 257 T; 0 U; 0 Other;

Query Match   84.8%;   Score 17.8; DB 6;   Length 1560;
Best Local Similarity   90.5%;   Pred. No. 87;
Matches   19; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 ATGACCACCAAGGCAGATGG 21
            ||||||||| ||||| |||
Db   1112 ATGACCACCACGAGGCACATGG 1132

RESULT 5
ABL91763
ID   ABL91763 standard; DNA; 1560 BP.
XX
AC   ABL91763;
XX
DT   28-MAY-2002 (first entry)
XX
DE   Human polynucleotide SEQ ID NO 106.
XX
KW   Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;
KW   Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide;
KW   cytostatic; virucide; protozoacide; antibacterial; ds.
XX
OS   Homo sapiens.
XX
PN   DE10100586-C1.
XX
PD   11-APR-2002.
XX
PF   09-JAN-2001; 2001DE-01000586.
XX
PR   09-JAN-2001; 2001DE-01000586.
XX
PA   (RIBO-) RIBOPHARMA AG.
XX
PI   Kreutzer R,   Limmer S,   Rost S,   Hadwiger P;
XX
DR   WPI; 2002-270454/32.
XX
PT   Inhibiting gene expression in cells, useful for e.g. treating tumors, by
PT   introducing double-stranded complementary oligoRNA having unpaired
PT   terminal bases.
XX
PS   Claim 13; Page 78; 104pp; German.
XX
CC   The invention relates to a method for inhibiting expression of a target
CC   gene (ABL91658-ABL91797) in a cell by introducing at least one
CC   oligoribonucleotide that has a double-stranded structure consisting of at
CC   most 49 sequential nucleotide pairs, with at least part of one strand
CC   complementary with the target gene and has at least one end a single-
CC   stranded segment of 1-4 nt. The method provides oligoribonucleotides for
CC   antisense inhibition of gene expression useful e.g. for treating tumours
CC   but the oligoribonucleotides may also be directed against genes present
CC   in pathogens (e.g. Plasmodium or viruses/viroids, pathogenic on humans,
CC   animals or plants) or against cytokine, Id, developmental or prion genes.
CC   The method provides more effective inhibition of gene expression than use
CC   of known oligonucleotides, probably because the unpaired overhang
CC   increases stability and thus intracellular concentration
XX
SQ   Sequence 1560 BP; 275 A; 549 C; 479 G; 257 T; 0 U; 0 Other;

Query Match   84.8%;   Score 17.8; DB 6;   Length 1560;
Best Local Similarity   90.5%;   Pred. No. 87;
Matches   19; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 ATGACCACCAAGGCAGATGG 21
            ||||||||| ||||| |||
Db   1112 ATGACCACCACGAGGCACATGG 1132

RESULT 6
ABV78191
ID   ABV78191 standard; DNA; 1818 BP.
XX
AC   ABV78191;
XX
DT   15-NOV-2002 (first entry)
XX
DE   Human MT4MMP DNA SEQ ID NO 75.
XX
KW   RNA inhibition; dsRNAi; gene expression inhibitor; oncogene; cytostatic;

KW   virucide; protozoacide; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200255693-A2.
XX
PD   18-JUL-2002.
XX
PF   09-JAN-2002; 2002WO-EP000152.
XX
PR   09-JAN-2001; 2001DE-01000586.
PR   26-OCT-2001; 2001DE-01055280.
PR   29-NOV-2001; 2001DE-01058411.
PR   07-DEC-2001; 2001DE-01060151.
XX
PA   (RIBO-) RIBOPHARMA AG.
XX
PI   Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
DR   WPI; 2002-590671/63.
XX
PT   Inhibiting expression of target gene, useful e.g. for inhibiting
PT   oncogenes, by administering double-stranded RNA complementary to the
PT   target and having an overhang.
XX
PS   Claim 10; Page 153-154; 203pp; German.
XX
CC   The invention relates to inhibiting expression of a target gene (I) in a
CC   cell by introducing an inhibitory RNA (dsRNAI) having a double-stranded
CC   structure of at most 49 consecutive bases. At least part of one strand
CC   (as1) of dsRNAi is complementary to (I) and at least one end of dsRNA1
CC   has an overhang of 1-4 nucleotides. The method is used to inhibit the
CC   expression of a wide range of genes, e.g. oncogenes, cytokine genes etc.
CC   in humans, also genes in Plasmodium or in viruses or viroids that are
CC   pathogenic for humans, animals or plants. Introducing an overhang into
CC   dsRNA greatly increases effectiveness for inhibiting gene expression,
CC   both in vivo and in vitro and also increases stability and thus the
CC   effective concentration inside the cell. The present sequence is that of
CC   a gene related to the invention
XX
SQ   Sequence 1818 BP; 310 A; 649 C; 571 G; 288 T; 0 U; 0 Other;

Query Match   84.8%; Score 17.8; DB 6; Length 1818;
Best Local Similarity   90.5%; Pred. No. 89;
Matches   19; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 ATGACCACCAAGGCAGATGG 21
         |||||||||| ||||| ||||
Db 1370 ATGACCACCGAGGCACATGG 1390

RESULT 7
ABZ35767
ID   ABZ35767 standard; DNA; 1818 BP.
XX
AC   ABZ35767;
XX
DT   07-FEB-2003 (first entry)
XX
DE   Human MT4MMP polynucleotide SEQ ID NO 75.
XX
KW   Double stranded RNA; dsRNA; RNAi; RNA inhibition; cytostatic; virucide;
KW   protozoacide; gene expression; antisense; tumour; infection; Plasmodium;
KW   virus; viroid; anti-GFP; human; HIV; human immunodeficiency virus;
KW   Hepatitis C virus; human papilloma virus; gene; ds.
XX
OS   Homo sapiens.
XX
PN   DE10100588-A1.
XX
PD   18-JUL-2002.
XX
PF   09-JAN-2001; 2001DE-01000588.
XX
PR   09-JAN-2001; 2001DE-01000588.
XX
PA   (RIBO-) RIBOPHARMA AG.
XX
PI   Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
DR   WPI; 2002-683450/74.
XX
PT   Inhibiting expression of target genes, useful e.g. for treating tumors,
PT   by introducing into cells two double-stranded RNAs that are complementary
PT   to the target.
XX
PS   Claim 13; Page 50; 100pp; German.
XX
CC   The invention relates to inhibiting expression of a target gene in a cell
CC   by introducing at least two oligoribonucleotides (dsRNAI and II), both
CC   with a double-stranded (ds) structure of at most 49 sequential nucleotide
CC   pairs. At least part of one strand (S1, S2) of the ds structures in each
CC   of dsRNAi and II are complementary to regions in the target gene. The
CC   method uses antisense inhibition of gene expression using double stranded
CC   RNA inhibition (RNAi). The method is particularly used to treat tumours
CC   or infections, especially by Plasmodium or viruses/viroids (pathogenic on
CC   humans, animals or plants). The method provides more effective inhibition
CC   of expression than known methods using a single dsRNA, even at very low
CC   concentrations. When dsRNA has at least one unpaired nucleotide at the
CC   end, stability (and thus effective concentration in the cell) is improved
CC   and efficiency can be increased further by pretreating the cells with
CC   interferon. The present sequence is that of a target DNA of the invention
XX
SQ   Sequence 1818 BP; 310 A; 649 C; 571 G; 288 T; 0 U; 0 Other;

Query Match   84.8%; Score 17.8; DB 6; Length 1818;
Best Local Similarity   90.5%; Pred. No. 89;
Matches   19; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 ATGACCACCAAGGCAGATGG 21
         |||||||||| ||||| ||||
Db 1370 ATGACCACCGAGGCACATGG 1390

RESULT 8
ABX10010
ID   ABX10010 standard; DNA; 1818 BP.
XX
AC   ABX10010;
XX
DT   23-JAN-2003 (first entry)
XX
DE   Human MT4MMP DNA fragment SEQ ID 75.
XX
KW   Oligoribonucleotide; interferon; oncogene; cytokine; Id; developmental;
KW   prion; inhibition; human; ds.
XX
OS   Homo sapiens.
XX
PN   DE10100587-C1.
XX
PD   21-NOV-2002.
XX
PF   09-JAN-2001; 2001DE-01000587.
XX
PR   09-JAN-2001; 2001DE-01000587.
XX
PA   (RIBO-) RIBOPHARMA AG.
XX
PI   Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
DR   WPI; 2002-742209/81.
XX
PT   Inhibiting expression of target genes, e.g. oncogenes, in cells, by
PT   introduction of complementary double-stranded oligoribonucleotide, after
PT   treating the cell with interferon.

XX
PS Disclosure; Page 55; 98pp; German.
XX
CC This invention describes a novel method for inhibiting expression of a
CC target gene by introducing into the cell that contains the target gene at
CC least one oligoribonucleotide (dsRNAi) that has a double-stranded (ds)
CC structure of not more than 49 consecutive nucleotides (nt), where at
CC least a segment of one strand of the ds structure is complementary with
CC the target gene and the cells are treated with interferon before
CC introduction of dsRNAi. The method is used to inhibit expression of
CC target genes, particularly oncogenes, cytokine genes, Id (not defined)
CC protein genes; developmental or prion genes, or genes expressed in
CC pathogenic organisms (particularly plasmodia) or in viruses or viroids
CC (pathogenic in humans, animals or plants). Treating the cells with
CC interferon greatly increases the extent to which dsRNA can inhibit
CC expression of the target genes, and the effect is even greater when dsRNA
CC are modified to increase their stability. ABX09936-ABX10075 represent
CC gene fragments used to illustrate the method of the invention
XX
SQ Sequence 1818 BP; 310 A; 649 C; 571 G; 288 T; 0 U; 0 Other;

Query Match   84.8%; Score 17.8; DB 6; Length 1818;
Best Local Similarity   90.5%; Pred. No. 89;
Matches 19; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 ATGACCACCACCAAGGCAGATGG 21
          |||||||||| |||| |||| |||
Db   1370 ATGACCACCACGAGGCACATGG 1390

RESULT 9
ABL91732
ID ABL91732 standard; DNA; 1818 BP.
XX
AC ABL91732;
XX
DT 28-MAY-2002 (first entry)
XX
DE Human polynucleotide SEQ ID NO 75.
XX
KW Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;
KW Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide;
KW cytostatic; virucide; protozoacide; antibacterial; ds.
XX
OS Homo sapiens.
XX
PN DE10100586-C1.
XX
PD 11-APR-2002.
XX
PF 09-JAN-2001; 2001DE-01000586.
XX
PR 09-JAN-2001; 2001DE-01000586.
XX
PA (RIBO-) RIBOPHARMA AG.
XX
PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
DR WPI; 2002-270454/32.
XX
PT Inhibiting gene expression in cells, useful for e.g. treating tumors, by
PT introducing double-stranded complementary oligoRNA having unpaired
PT terminal bases.
XX
PS Claim 13; Page 53-54; 104pp; German.
XX
CC The invention relates to a method for inhibiting expression of a target
CC gene (ABL91658-ABL91797) in a cell by introducing at least one
CC oligoribonucleotide that has a double-stranded structure consisting of at
CC most 49 sequential nucleotide pairs, with at least part of one strand
CC complementary with the target gene and has at least one end a single-
CC stranded segment of 1-4 nt. The method provides oligoribonucleotides for
CC antisense inhibition of gene expression useful e.g. for treating tumours

XX
CC but the oligoribonucleotides may also be directed against genes present
CC in pathogens (e.g. Plasmodium or viruses/viroids, pathogenic on humans,
CC animals or plants) or against cytokine, Id, developmental or prion genes.
CC The method provides more effective inhibition of gene expression than use
CC of known oligonucleotides, probably because the unpaired overhang
CC increases stability and thus intracellular concentration
XX
SQ Sequence 1818 BP; 310 A; 649 C; 571 G; 288 T; 0 U; 0 Other;

Query Match   84.8%; Score 17.8; DB 6; Length 1818;
Best Local Similarity   90.5%; Pred. No. 89;
Matches 19; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 ATGACCACCACCAAGGCAGATGG 21
          |||||||||| |||| |||| |||
Db   1370 ATGACCACCACGAGGCACATGG 1390

RESULT 10
AAA13389
ID AAA13389 standard; DNA; 2295 BP.
XX
AC AAA13389;
XX
DT 28-JUL-2000 (first entry)
XX
DE Human MT4-MMP matrixmetalloprotease nucleotide sequence.
XX
KW Matrix metalloprotease; MT4-MMP; treatment; arthrosis deformans;
KW chronic rheumatoid arthritis; asthma; autoimmune disease; hepatitis;
KW atrophic dermatitis; psoriasis; contact dermatitis; hair loss; nephritis;
KW ischaemic disease; pancreatitis, arteriosclerosis; leukaemia; tumour;
KW wound; corneal ulcer; tissue damage; inflammation; cerebral apoplexia;
KW Alzheimer's disease; dementia; multiple sclerosis; Parkinson's disease;
KW gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200018900-A1.
XX
PD 06-APR-2000.
XX
PF 29-SEP-1999;   99WO-JP005349.
XX
PR 29-SEP-1998;   98JP-00276258.
PR 29-SEP-1998;   98JP-00291505.
XX
PA (SEIK/) SEIKI M.
XX
PI Seiki M;
XX
DR WPI; 2000-293141/25.
XX
PT Physiologically-active transmembrane matrix metalloprotease polypeptide,
PT useful in screening inhibitors and activators for treating e.g. arthrosis
PT deformans, asthma and cancers.
XX
PS Example 1; Page 99-102; 113pp; Japanese.
XX
CC This sequence represents the human matrix metalloprotease (MT4-MMP 2)
CC nucleotide sequence. The invention relates to a mammalian transmembrane
CC MMP polypeptide, its variants and the DNA sequences encoding them. Also
CC included in the invention is a vector comprising the MMP DNA molecules,
CC and cells transformed using the vector. The protein, DNA and antibodies
CC directed against the polypeptide are useful in screening inhibitors and
CC activators for use in treating arthrosis deformans, chronic rheumatoid
CC arthritis, asthma, autoimmune diseases, atrophic dermatitis, psoriasis,
CC contact dermatitis, hair loss, ischaemic diseases, immune reaction
CC accompanying organ transplant, hepatitis, nephritis, pancreatitis,
CC arteriosclerosis, leukaemia, malignant tumours, wounds, corneal ulcers,
CC tissue damage or inflammation accompanying leukocytic infiltration,
CC together with brain disorders during cerebral apoplexia, Alzheimer's
CC disease, dementia, multiple sclerosis, Parkinson's disease or brain

CC   tumours. The DNA can also be incorporated into a suitable vector for use
CC   in gene therapy
XX
SQ   Sequence 2295 BP; 401 A; 807 C; 729 G; 358 T; 0 U; 0 Other;

Query Match   84.8%;   Score 17.8; DB 3; Length 2295;
Best Local Similarity   90.5%;   Pred. No. 92;
Matches   19; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 ATGACCACCAAGGCAGATGG 21
            |||||||||| ||||| ||||
Db   1341 ATGACCACACGAGGCACATGG 1361

RESULT 11
AAA14351
ID   AAA14351 standard; DNA; 2295 BP.
XX
AC   AAA14351;
XX
DT   15-AUG-2000 (first entry)
XX
DE   Human MT4-MMP(2) gene fragment, used as a probe.
XX
KW   Matrix metalloprotease; MT4-MMP(2); transmembrane; human; antibody;
KW   drug screening; diagnosis; arthrosis deformans; rheumatoid arthritis;
KW   asthma; autoimmune disease; diagnosis; atopic dermatitis; skin disorder;
KW   ischaemic disorder; arteriosclerosis; cancer; tissue damage;
KW   inflammatory disorder; neurological disorder; probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200018805-A1.
XX
PD   06-APR-2000.
XX
PF   29-SEP-1999;   99WO-JP005350.
XX
PR   29-SEP-1998;   98JP-00291501.
PR   29-SEP-1998;   98JP-00291503.
XX
PA   (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI   Hanai N, Furuya A;
XX
DR   WPI; 2000-293115/25.
XX
PT   Antibody immunospecific for transmembrane matrix metalloprotease
PT   polypeptide, useful in screening of inhibitors and activators as drugs
PT   for, and for diagnosis of e.g. arthrosis deformans and asthma.
XX
PS   Example 1; Page 134-137; 152pp; Japanese.
XX
CC   The invention relates to antibodies against the novel murine and human
CC   transmembrane matrix metalloproteases MT4-MMP(2) (AAY90502 and AAY90503)
CC   and MT5-MMP (AAY90504 and AAY90505). The antibodies are useful for the
CC   diagnosis and screening of inhibitors and activators such as arthrosis
CC   or preventing a wide variety of medical conditions, such as arthrosis
CC   deformans, rheumatoid arthritis, asthma, autoimmune diseases and atopic
CC   dermatitis. The anti-MT5-MMP antibodies may additionally be used to
CC   diagnose and screen therapeutic or prophylactic agents for conditions
CC   affecting the brain, such as cerebral stroke and Alzheimer's disease. The
CC   antibodies of the invention may also be used for the prevention,
CC   diagnosis and treatment of psoriasis, contact dermatitis, hair loss,
CC   ischaemic diseases, immune reaction accompanying organ transplant,
CC   hepatitis, nephritis, pancreatitis, arteriosclerosis, leukaemia,
CC   malignant tumours, wounds, corneal ulcers, tissue damage or inflammation
CC   accompanying leukocytic infiltration, dementia, multiple sclerosis,
CC   Parkinson's disease or brain tumour. The present sequence represents a
CC   fragment of the human MT4-MMP(2) gene used as a probe in an
CC   exemplification of the invention to screen a murine cDNA library for a
CC   murine MT4-MMP(2) cDNA clone
XX

SQ   Sequence 2295 BP; 401 A; 807 C; 729 G; 358 T; 0 U; 0 Other;

Query Match   84.8%;   Score 17.8; DB 3; Length 2295;
Best Local Similarity   90.5%;   Pred. No. 92;
Matches   19; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 ATGACCACCAAGGCAGATGG 21
            |||||||||| ||||| ||||
Db   1341 ATGACCACACGAGGCACATGG 1361

RESULT 12
AAH28234
ID   AAH28234 standard; cDNA; 2306 BP.
XX
AC   AAH28234;
XX
DT   05-SEP-2001 (first entry)
XX
DE   Nucleotide sequence of matrix metalloproteinase-17.
XX
KW   Growth factor; protein inhibitor; protease; damaged tissue;
KW   platelet-derived growth factor; PDGF; fibroblast growth factor; FGF;
KW   connective tissue derived growth factor; CTGF; chrysalin; VEGF;
KW   keratinocyte-derived growth factor; KGF; epidermal growth factor; EGF;
KW   transforming growth factor-beta; TGF-beta; matrix metalloproteinase; MMP;
KW   granulocyte macrophage colony stimulating factor; GM-CSF; uPA;
KW   vascular endothelial growth factor; urokinase plasminogen activator;
KW   dermal ulcer; wound; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   233..1792
FT   /*tag= a
FT   /product= "MMP-17"
XX
PN   WO200149309-A2.
XX
PD   12-JUL-2001.
XX
PF   21-DEC-2000;   2000WO-IB001935.
XX
PR   29-DEC-1999;   99GB-00030768.
XX
PA   (PFIZ) PFIZER LTD.
PA   (PFIZ) PFIZER INC.
XX
PI   Davies MJ, Huggins JP, Mcintosh FS,   Occleston NL;
XX
DR   WPI; 2001-418351/44.
DR   P-PSDB; AAB84619.
XX
PT   Composition for the treatment of damaged tissue i.e. chronic wounds and
PT   dermal ulcers comprises an inhibitor agent i.e. a protease and a growth
PT   factor.
XX
PS   Disclosure; Page 567; 572pp; English.
XX
CC   The specification describes a pharmaceutical composition, comprising a
CC   growth factor, an inhibitor agent, i.e. a protease. The inhibitor agent
CC   inhibits the action of at least one specific adverse protein, i.e. a
CC   protease, that is upregulated in a damaged tissue such as a wound
CC   environment. Growth factors which are included in the composition of the
CC   invention are platelet-derived growth factor (PDGF), fibroblast growth
CC   factor (FGF), connective tissue derived growth factor (CTGF),
CC   keratinocyte-derived growth factor (KGF), transforming growth factor-beta
CC   (TGF-beta), granulocyte macrophage colony stimulating factor (GM-CSF),
CC   epidermal growth factor (EGF), vascular endothelial growth factor (VEGF),
CC   and chrysalin. Inhibitors which are included in the composition of the
CC   invention include inhibitors of urokinase-type plasminogen activator
CC   (uPA) and matrix metalloproteinase (MMP). The composition is useful for
CC   the treatment of chronic damaged tissue, i.e. wounds and dermal ulcers.

CC   The present sequence encodes a human MMP-17, and is used to produce the
CC   composition of the invention
XX
SQ   Sequence 2306 BP; 402 A; 810 C; 735 G; 359 T; 0 U; 0 Other;

Query Match   84.8%;   Score 17.8;   DB 4;   Length 2306;
Best Local Similarity   90.5%;   Pred. No. 92;
Matches   19;   Conservative   0;   Mismatches   2;   Indels   0;   Gaps   0;

Qy   1 ATGACCACACAAGGCAGATGG 21
            ||||||||||| ||||| ||||
Db   1344 ATGACCACCACGAGGCACATGG 1364

RESULT 13
ADB84832
ID   ADB84832 standard; DNA; 2306 BP.
XX
AC   ADB84832;
XX
DT   29-JAN-2004 (first entry)
XX
DE   Farnesyl transferase inhibitor modulated leukemia associated gene #51.
XX
KW   ss; cytostatic; farnesyl transferase inhibitor; gene expression;
KW   quinolinone; leukemia; cancer.
XX
OS   Homo sapiens.
XX
PN   WO2003038129-A2.
XX
PD   08-MAY-2003.
XX
PF   30-OCT-2002; 2002WO-US034784.
XX
PR   30-OCT-2001; 2001US-0338997P.
PR   30-OCT-2001; 2001US-0340081P.
PR   30-OCT-2001; 2001US-0340938P.
PR   30-OCT-2001; 2001US-0341012P.
XX
PA   (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.
XX
PI   Raponi M;
XX
DR   WPI; 2003-513497/48.
XX
PT   Determining whether a patient will respond to treatment with a farnesyl
PT   transferase inhibitor, by analyzing the expression of gene that is
PT   differentially modulated in the presence of the inhibitor.
XX
PS   Disclosure; SEQ ID NO 51; 346pp; English.
XX
CC   The invention relates to a method of determining whether a patient will
CC   respond to treatment with a farnesyl transferase inhibitor (FTI), by
CC   analyzing the expression of gene that is differentially modulated in the
CC   presence of an FTI. The method is useful for determining whether a
CC   patient will respond to treatment with a FTI such as (B)-6-[amino(4-
CC   chlorophenyl)(1-methyl-1H-imidazol-5-yl)methyl]-4-(3-chlorophenyl)-1-
CC   methyl-2-(1H)quinolinone, monitoring the therapy of a patient, treating a
CC   patient with leukemia with FTI if the analysis indicates that the patient
CC   will respond. This sequence corresponds to a gene whose expression may be
CC   modulated in the presence of FTI.
XX
SQ   Sequence 2306 BP; 402 A; 810 C; 735 G; 359 T; 0 U; 0 Other;

Query Match   84.8%;   Score 17.8;   DB 10;   Length 2306;
Best Local Similarity   90.5%;   Pred. No. 92;
Matches   19;   Conservative   0;   Mismatches   2;   Indels   0;   Gaps   0;

Qy   1 ATGACCACACAAGGCAGATGG 21
            ||||||||||| ||||| ||||
Db   1344 ATGACCACCACGAGGCACATGG 1364

RESULT 14
AAA14340
ID   AAA14340 standard; cDNA; 2423 BP.
XX
AC   AAA14340;
XX
DT   15-AUG-2000 (first entry)
XX
DE   cDNA encoding human transmembrane matrix metalloprotease MT4-MMP(2).
XX
KW   Matrix metalloprotease; MT4-MMP(2); transmembrane; human; antibody;
KW   drug screening; diagnosis; arthrosis deformans; rheumatoid arthritis;
KW   asthma; autoimmune disease; atopic dermatitis; skin disorder;
KW   ischaemic disorder; arteriosclerosis; cancer; tissue damage;
KW   inflammatory disorder; neurological disorder; ss.
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   100..1917
FT   /tag= a
FT   /product= "Human transmembrane matrix metalloprotease MT4
FT   -MMP(2)"
XX
PN   WO200018805-A1.
XX
PD   06-APR-2000.
XX
PF   29-SEP-1999;   99WO-JP005350.
XX
PR   29-SEP-1998;   98JP-00291501.
PR   29-SEP-1998;   98JP-00291503.
XX
PA   (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI   Hanai N, Furuya A;
XX
DR   WPI; 2000-293115/25.
DR   P-PSDB; AAY90503.
XX
PT   Antibody immunospecific for transmembrane matrix metalloprotease
PT   polypeptide, useful in screening of inhibitors and activators as drugs
PT   for, and for diagnosis of e.g. arthrosis deformans and asthma.
XX
PS   Example 2; Page 102-108; 152pp; Japanese.
XX
CC   This sequence represents cDNA encoding a novel human transmembrane matrix
CC   metalloprotease MT4-MMP(2). MT4-MMP(2) has physiological activity
CC   different to that of MT4-MMP. The invention relates to antibodies against
CC   both murine and human MT4-MMP(2) (AAY90502 and AAY90503) and to
CC   antibodies against a novel murine and human transmembrane matrix
CC   metalloprotease MT5-MMP (AAY90504 and AAY90505). The antibodies are
CC   useful for the diagnosis and screening of inhibitors and activators
CC   useful for treating or preventing a wide variety of medical conditions,
CC   such as arthrosis deformans, rheumatoid arthritis, asthma, autoimmune
CC   diseases and atopic dermatitis. The anti-MT5-MMP antibodies may
CC   additionally be used to diagnose and screen therapeutic or prophylactic
CC   agents for conditions affecting the brain, such as cerebral stroke and
CC   Alzheimer's disease. The antibodies of the invention may also be used for
CC   the prevention, diagnosis and treatment of psoriasis, contact dermatitis,
CC   hair loss, ischaemic diseases, immune reaction accompanying organ
CC   transplant, hepatitis, nephritis, pancreatitis, arteriosclerosis,
CC   leukaemia, malignant tumours, wounds, corneal ulcers, tissue damage or
CC   inflammation accompanying leukocytic infiltration, dementia, multiple
CC   sclerosis, Parkinson's disease or brain tumour
XX
SQ   Sequence 2423 BP; 404 A; 871 C; 794 G; 354 T; 0 U; 0 Other;

Query Match   84.8%;   Score 17.8;   DB 3;   Length 2423;
Best Local Similarity   90.5%;   Pred. No. 93;
Matches   19;   Conservative   0;   Mismatches   2;   Indels   0;   Gaps   0;

Qy   1 ATGACCACACAAGGCAGATGG 21
            ||||||||||| ||||| ||||
Db   1344 ATGACCACCACGAGGCACATGG 1364

```
Qy          1 ATGACCACCAAGGCAGATGG 21
              ||||||||||| ||||| ||||
Db       1469 ATGACCACCACGCAGGCACATGG 1489
```

RESULT 15
ABZ20920
ID   ABZ20920 standard; cDNA; 2423 BP.
AC   ABZ20920;
XX
DT   26-MAR-2003 (first entry)
XX
DE   MT4-MMP catalytic domain antibody related coding sequence.
XX
KW   MT4-MMP; catalytic domain; antibody; inflammation; cancer; gene;
KW   membrane type-matrix metalloproteinase; antiinflammatory; antirheumatoid;
KW   antiarthritic; rheumatoid arthritis; cytostatic; human; ss.
XX
OS   Homo sapiens.
XX
PN   WO2002101046-A1.
XX
PD   19-DEC-2002.
XX
PF   11-JUN-2002; 2002WO-JP005788.
XX
PR   11-JUN-2001; 2001JP-00176256.
XX
PA   (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI   Miki I, Ohta S, Shitara K, Furuya A;
XX
DR   WPI; 2003-148808/14.
DR   P-PSDB; AAO19713.
XX
PT   Monoclonal antibody specifically binding to natural or solubilized MT4-
PT   MMP, applicable in diagnosis and remedies for MT4-MMP participated
PT   diseases e.g. inflammations and cancer particularly rheumatoid arthritis.
XX
PS   Example 1; Page 50-54; 63pp; Japanese.
XX
CC   The present invention relates to a monoclonal antibody which binds
CC   specifically to the MT4-MMP (membrane type-matrix metalloproteinase)
CC   catalytic domain. The antibody is applicable in diagnosis and remedies
CC   for MT4-MMP participated diseases e.g. inflammations and cancer
CC   particularly rheumatoid arthritis. The present sequence is a human coding
CC   sequence shown in the exemplification of the invention
XX
SQ   Sequence 2423 BP; 404 A; 871 C; 794 G; 354 T; 0 U; 0 Other;

Query Match   84.8%; Score 17.8; DB 8; Length 2423;
Best Local Similarity   90.5%; Pred. No. 93;
Matches   19; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 ATGACCACCAAGGCAGATGG 21
              ||||||||||| ||||| ||||
Db   1469 ATGACCACCACGCAGGCACATGG 1489
```

Search completed: May 15, 2005, 11:25:18
Job time : 17.1953 secs

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    May 15, 2005, 11:11:51 ; Search time 4.6102 Seconds
                                     (without alignments)
                              7453.434 Million cell updates/sec

Title:          US-10-764-075-2
Perfect score:  21
Sequence:       1 atgaccacacaaggcagatgg 21

Scoring table:  IDENTITY NUC
                Gapop 10.0 , Gapext 1.0

Searched:       1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters:    2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :     Issued Patents NA:*
               1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
               2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
               3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
               4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
               5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
               6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

       Pred. No. is the number of results predicted by chance to have a
       score greater than or equal to the score of the result being printed,
       and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.		Score	% Query Match	Match Length	DB	ID	Description
c	1	16.8	80.0	224	4	US-09-313-294A-7452	Sequence 7452, Ap
	2	16.8	80.0	2354	4	US-09-810-671-1	Sequence 1, Appli
	3	16.8	80.0	2354	4	US-10-109-854-1	Sequence 1, Appli
	4	16.8	80.0	2354	4	US-10-339-656-1	Sequence 1, Appli
	5	16.8	80.0	2446	2	US-09-016-000-9	Sequence 9, Appli
c	6	16.8	80.0	9769	3	US-08-961-527-30	Sequence 30, Appli
	7	16.8	80.0	21234	3	US-09-810-671-3	Sequence 3, Appli
	8	16.8	80.0	21234	4	US-10-109-854-3	Sequence 3, Appli
	9	16.8	80.0	21234	4	US-10-339-656-3	Sequence 3, Appli
c	10	16.8	80.0	100836	4	US-09-949-016-12871	Sequence 12871, A
c	11	16.8	80.0	100837	4	US-09-949-016-17063	Sequence 17063, A
c	12	16.8	80.0	149543	4	US-09-949-016-15947	Sequence 15947, A
c	13	16.8	80.0	248968	4	US-09-949-016-12614	Sequence 12614, A
c	14	16.8	80.0	250958	4	US-09-949-016-16061	Sequence 16061, A
c	15	16.2	77.1	601	4	US-09-949-016-169568	Sequence 169568,
	16	16.2	77.1	884	4	US-09-790-045-11	Sequence 11, Appl
	17	16.2	77.1	884	4	US-10-222-577-11	Sequence 11, Appl
	18	16.2	77.1	884	4	US-10-222-578-11	Sequence 11, Appl
	19	16.2	77.1	1200	4	US-09-598-401C-36	Sequence 36, Appl
	20	16.2	77.1	17000	4	US-09-679-299A-18	Sequence 18, Appl
	21	16.2	77.1	26103	4	US-09-949-016-16841	Sequence 16841, A
c	22	16.2	77.1	160759	4	US-09-949-016-16514	Sequence 16514, A
c	23	16.2	77.1	235064	4	US-09-949-016-15390	Sequence 15390, A
	24	16.2	77.1	421491	4	US-09-949-016-12805	Sequence 12805, A
c	25	16.2	77.1	421494	4	US-09-949-016-14060	Sequence 14060, A
	26	15.8	75.2	471	4	US-09-513-999C-992	Sequence 992, App
	27	15.8	75.2	537	4	US-09-513-999C-993	Sequence 993, App
c	28	15.8	75.2	601	4	US-09-949-016-91184	Sequence 91184, A
c	29	15.8	75.2	601	4	US-09-949-016-201853	Sequence 201853,
	30	15.8	75.2	1174	2	US-08-481-658B-39	Sequence 39, Appl
	31	15.8	75.2	1174	2	US-08-477-504A-39	Sequence 39, Appl
	32	15.8	75.2	1174	2	US-08-486-756A-39	Sequence 39, Appl
	33	15.8	75.2	1174	2	US-08-485-862B-39	Sequence 39, Appl
	34	15.8	75.2	1174	3	US-08-787-739-39	Sequence 39, Appl
	35	15.8	75.2	1174	3	US-08-487-077A-39	Sequence 39, Appl
	36	15.8	75.2	1174	3	US-08-485-863A-39	Sequence 39, Appl
	37	15.8	75.2	1174	3	US-08-485-049D-39	Sequence 39, Appl
	38	15.8	75.2	1174	3	US-09-178-115-39	Sequence 39, Appl
	39	15.8	75.2	1174	4	US-09-177-776-39	Sequence 39, Appl
	40	15.8	75.2	1174	4	US-09-772-719B-39	Sequence 39, Appl
	41	15.8	75.2	1505	4	US-09-620-312D-544	Sequence 544, App
	42	15.8	75.2	4079	4	US-09-016-434-1173	Sequence 1173, Ap
	43	15.8	75.2	4425	4	US-09-949-016-5564	Sequence 5564, Ap
	44	15.8	75.2	4426	3	US-09-658-687A-3	Sequence 3, Appli
	45	15.8	75.2	4426	4	US-09-949-016-843	Sequence 843, App

ALIGNMENTS

RESULT 1
US-09-313-294A-7452/c
```
; Sequence 7452, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 7452
; LENGTH: 224
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700381752H1
US-09-313-294A-7452

Query Match          80.0%; Score 16.8; DB 4; Length 224;
Best Local Similarity 90.0%; Pred. No. 74;
Matches  18; Conservative  0; Mismatches  2; Indels  0; Gaps  0;

Qy          2 TGACCACAAGGCAGATGG 21
              || |||||||| |||||||
Db         39 TGTCCACACAAGGCTGATGG 20
```

RESULT 2
US-09-810-671-1
```
; Sequence 1, Application US/09810671
; Patent No. 6455291
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758
; CURRENT APPLICATION NUMBER: US/09/810,671
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2354
; TYPE: DNA
; ORGANISM: Human
```

US-09-810-671-1

Query Match   80.0%;   Score 16.8; DB 3;   Length 2354;
Best Local Similarity   90.0%;   Pred. No. 1e+02;
Matches   18; Conservative   0; Mismatches   2; Indels   2; Gaps   0;

```
Qy          2 TGACCACACAAGGCAGATGG 21
              ||||||| |||||||||||
Db        704 TGACCACATCAGGCAGATGG 723
```

RESULT 3
US-10-109-854-1
; Sequence 1, Application US/10109854
; Patent No. 6630337
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758DIV
; CURRENT APPLICATION NUMBER: US/10/109,854
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2354
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-109-854-1

Query Match   80.0%;   Score 16.8; DB 4;   Length 2354;
Best Local Similarity   90.0%;   Pred. No. 1e+02;
Matches   18; Conservative   0; Mismatches   2; Indels   2; Gaps   0;

```
Qy          2 TGACCACACAAGGCAGATGG 21
              ||||||| |||||||||||
Db        704 TGACCACATCAGGCAGATGG 723
```

RESULT 4
US-10-339-656-1
; Sequence 1, Application US/10339656
; Patent No. 6733978
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758DIV2
; CURRENT APPLICATION NUMBER: US/10/339,656
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/109,854
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2354
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-339-656-1

Query Match   80.0%;   Score 16.8; DB 4;   Length 2354;
Best Local Similarity   90.0%;   Pred. No. 1e+02;
Matches   18; Conservative   0; Mismatches   2; Indels   2; Gaps   0;

```
Qy          2 TGACCACACAAGGCAGATGG 21
              ||||||| |||||||||||
Db        704 TGACCACATCAGGCAGATGG 723
```

RESULT 5
US-09-016-000-9
; Sequence 9, Application US/09016000
; Patent No. 5962232
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Bandman, Olga
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; TITLE OF INVENTION: PROTEIN KINASE MOLECULES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Incyte Pharmaceuticals, Inc.
;   STREET: 3174 Porter Drive
;   CITY: Palo Alto
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94304
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/016,000
;   FILING DATE: HEREWITH
;   CLASSIFICATION:
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
; ATTORNEY/AGENT INFORMATION:
;   NAME: Billings, Lucy J
;   REGISTRATION NUMBER: 36,749
;   REFERENCE/DOCKET NUMBER: PF-0465 US
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 650-855-0555
;   TELEFAX: 650-845-4166
;   TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 2446 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; IMMEDIATE SOURCE:
;   LIBRARY: NEUTFMT01
;   CLONE: 339963
US-09-016-000-9

Query Match   80.0%;   Score 16.8; DB 2;   Length 2446;
Best Local Similarity   90.0%;   Pred. No. 1e+02;
Matches   18; Conservative   0; Mismatches   2; Indels   2; Gaps   0;

```
Qy          2 TGACCACACAAGGCAGATGG 21
              ||||||| |||||||||||
Db        817 TGACCACATCAGGCAGATGG 836
```

RESULT 6
US-08-961-527-30/c
; Sequence 30, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch

TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 9769 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-30

Query Match   80.0%;   Score 16.8; DB 3;   Length 9769;
Best Local Similarity   90.0%;   Pred. No. 1.3e+02;
Matches   18;   Conservative   0; Mismatches   2;   Indels   0;   Gaps   0;

Qy   2 TGACCACACAAGGCAGATGG 21
          |||||||| |||||||||
Db   3808 TGACCAAACAAGGCAGTTGG 3789

RESULT 7
US-09-810-671-3
; Sequence 3, Application US/09810671
; Patent No. 6455291
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758
; CURRENT APPLICATION NUMBER: US/09/810,671
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 21234
; TYPE: DNA
; ORGANISM: Human
US-09-810-671-3

Query Match   80.0%;   Score 16.8; DB 3;   Length 21234;
Best Local Similarity   90.0%;   Pred. No. 1.4e+02;
Matches   18;   Conservative   0; Mismatches   2;   Indels   0;   Gaps   0;

Qy   2 TGACCACACAAGGCAGATGG 21
          |||||||| ||||||||
Db   8379 TGACCACATCAGGCAGATGG 8398

RESULT 8
US-10-109-854-3
; Sequence 3, Application US/10109854
; Patent No. 6630337
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758DIV
; CURRENT APPLICATION NUMBER: US/10/109,854
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 21234
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-109-854-3

Query Match   80.0%;   Score 16.8; DB 4;   Length 21234;
Best Local Similarity   90.0%;   Pred. No. 1.4e+02;
Matches   18;   Conservative   0; Mismatches   2;   Indels   0;   Gaps   0;

Qy   2 TGACCACACAAGGCAGATGG 21
          |||||||| ||||||||
Db   8379 TGACCACATCAGGCAGATGG 8398

RESULT 9
US-10-339-656-3
; Sequence 3, Application US/10339656
; Patent No. 6733978
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758DIV2
; CURRENT APPLICATION NUMBER: US/10/339,656
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/109,854
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 21234
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-339-656-3

Query Match   80.0%;   Score 16.8; DB 4;   Length 21234;
Best Local Similarity   90.0%;   Pred. No. 1.4e+02;
Matches   18;   Conservative   0; Mismatches   2;   Indels   0;   Gaps   0;

Qy   2 TGACCACACAAGGCAGATGG 21
          |||||||| ||||||||
Db   8379 TGACCACATCAGGCAGATGG 8398

RESULT 10
US-09-949-016-12871/c
; Sequence 12871, Application US/0994901 6
; Patent No. 6812339

```
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12871
; LENGTH: 100836
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(100836)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12871

Query Match     80.0%;  Score 16.8;  DB 4;  Length 100836;
Best Local Similarity   90.0%;  Pred. No. 1.7e+02;
Matches  18;  Conservative   0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy     2 TGACCACACAAGGCAGATGG 21
          |||||||||||||| |||||
Db 68341 TGACCACACAAGGTTGATGG 68322

RESULT 11
US-09-949-016-17063/c
; Sequence 17063, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17063
; LENGTH: 100837
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(100837)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17063

Query Match     80.0%;  Score 16.8;  DB 4;  Length 100837;
Best Local Similarity   90.0%;  Pred. No. 1.7e+02;
Matches  18;  Conservative   0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy     2 TGACCACACAAGGCAGATGG 21
          |||||||||||||| |||||
Db 68341 TGACCACACAAGGTTGATGG 68322

RESULT 12
US-09-949-016-15947/c
; Sequence 15947, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15947
; LENGTH: 149543
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15947

Query Match     80.0%;  Score 16.8;  DB 4;  Length 149543;
Best Local Similarity   90.0%;  Pred. No. 1.9e+02;
Matches  18;  Conservative   0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy     2 TGACCACACAAGGCAGATGG 21
          |||||||||||||| |||||
Db 52555 TGACCACACAAGGTTGATGG 52536

RESULT 13
US-09-949-016-12614/c
; Sequence 12614, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12614
; LENGTH: 248968
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(248968)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12614

Query Match     80.0%;  Score 16.8;  DB 4;  Length 248968;
Best Local Similarity   90.0%;  Pred. No. 1.9e+02;
Matches  18;  Conservative   0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy     1 ATGACCACACAAGGCAGATG 20
          |||| ||||||||||||||||
Db 67410 ATGAAAACACAAGGCAGATG 67391

RESULT 14
US-09-949-016-16061/c
```

```
; Sequence 16061, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE: 2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16061
;  LENGTH: 250958
;  TYPE: DNA
;  ORGANISM: Human
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (1)...(250958)
;  OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16061

Query Match             80.0%;  Score 16.8;  DB 4;  Length 250958;
Best Local Similarity   90.0%;  Pred. No. 1.9e+02;
Matches   18;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy         1 ATGACCACACAAGGCAGATG 20
             |||| ||||||||||||||||
Db    117400 ATGAAAACACAAGGCCAGATG 117381


RESULT 15
US-09-949-016-169568/c
; Sequence 169568, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE: 2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 169568
;  LENGTH: 601
;  TYPE: DNA
;  ORGANISM: Human
US-09-949-016-169568

Query Match             77.1%;  Score 16.2;  DB 4;  Length 601;
Best Local Similarity   85.7%;  Pred. No. 1.7e+02;
Matches   18;  Conservative   0;  Mismatches   3;  Indels   0;  Gaps   0;

Qy       1 ATGACCACACAAGGCAGATGG 21
           ||||| |||| ||||| |||||
Db     455 ATGAACACACCAGGCACATGG 435


Search completed: May 15, 2005, 14:23:03
Job time : 8.6102 secs
```

THIS PAGE LEFT BLANK

OM nucleic - nucleic search, using sw model

```
            GenCore version 5.1.6
        Copyright (c) 1993 - 2005   Compugen Ltd.

Run on:     May 15, 2005, 12:53:36 ; Search time 15.5004 Seconds
                    (without alignments)
                    8291.679 Million cell updates/sec

Title:      US-10-764-075-2
Perfect score: 21
Sequence:   1 atgaccacacaaggcagatgg 21

Scoring table: IDENTITY_NUC
            Gapop 10.0 , Gapext 1.0

Searched:   5562332 seqs, 306019652 residues

Total number of hits satisfying chosen parameters:    11124664

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries
```

Database :   Published Applications NA:*

```
 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
 3: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
 4: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
 5: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
 6: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
 7: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
 8: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
 9: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

SUMMARIES

Result No.	Score	% Query Match	Match	Length	DB	ID	Description
1	21	100.0	21	21	18	US-10-764-075-2	Sequence 2, Appli
2	21	100.0	21	1136	18	US-10-764-075-1	Sequence 1, Appli
3	21	100.0	21	1620	19	US-10-699-936-16	Sequence 16, Appli
4	21	100.0	21	2304	19	US-10-699-936-7	Sequence 7, Appli
5	21	100.0	21	2304	19	US-10-699-936-11	Sequence 11, Appli
6	21	100.0	21	24774	19	US-10-889-447-3	Sequence 3, Appli
7	21	100.0	21	28920	19	US-10-889-447-5	Sequence 5, Appli
8	21	100.0	21	28920	19	US-10-889-447-6	Sequence 6, Appli
9	21	100.0	21	29291	19	US-10-889-447-4	Sequence 4, Appli
10	21	100.0	21	29430	19	US-10-889-447-7	Sequence 7, Appli
11	21	100.0	21	29727	18	US-10-839-729-15	Sequence 15, Appli
12	21	100.0	21	29727	18	US-10-827-757-1	Sequence 1, Appli
13	21	100.0	21	29727	19	US-10-889-447-8	Sequence 8, Appli
14	21	100.0	21	29727	19	US-10-699-936-1	Sequence 1, Appli
15	21	100.0	21	29736	18	US-10-839-729-17	Sequence 17, Appli
16	21	100.0	21	29736	19	US-10-889-447-9	Sequence 9, Appli
17	21	100.0	21	29736	18	US-10-699-936-3	Sequence 3, Appli
18	21	100.0	21	29742	19	US-10-839-729-16	Sequence 16, Appli
19	21	100.0	21	29742	19	US-10-808-187-15	Sequence 15, Appl
20	21	100.0	21	29742	19	US-10-808-187-16	Sequence 16, Appl
21	21	100.0	21	29742	19	US-10-808-187-240	Sequence 240, App
22	21	100.0	21	29742	19	US-10-808-187-737	Sequence 737, App
23	21	100.0	21	29742	19	US-10-808-187-1108	Sequence 1108, Ap
24	21	100.0	21	29742	19	US-10-808-187-1590	Sequence 1590, Ap
25	21	100.0	21	29742	19	US-10-808-187-1965	Sequence 1965, Ap
26	21	100.0	21	29742	19	US-10-889-447-10	Sequence 10, Appl
27	21	100.0	21	29751	18	US-10-839-729-14	Sequence 14, Appl
28	21	100.0	21	29751	19	US-10-856-529-1	Sequence 1, Appli
29	21	100.0	21	29751	19	US-10-626-879-67	Sequence 67, Appl
30	21	100.0	21	29751	19	US-10-889-447-1	Sequence 1, Appli
31	21	100.0	21	29751	19	US-10-889-447-2	Sequence 2, Appli
32	21	100.0	21	29751	19	US-10-699-936-2	Sequence 2, Appli
33	17.8	84.8		553	13	US-10-027-632-305207	Sequence 305207,
34	17.8	84.8		553	17	US-10-027-632-305207	Sequence 305207,
35	17.8	84.8		1560	18	US-10-384-339C-106	Sequence 106, App
36	17.8	84.8		1818	18	US-10-384-339C-75	Sequence 75, Appl
37	17.8	84.8		2295	16	US-10-406-209-17	Sequence 17, Appl
38	17.8	84.8		2306	16	US-10-131-985-50	Sequence 50, Appl
39	17.8	84.8		2306	18	US-10-283-975A-51	Sequence 51, Appl
40	17.8	84.8		2306	18	US-10-901-417-50	Sequence 50, Appl
41	17.8	84.8		2306	19	US-10-425-114-26875	Sequence 26875, A
42	17.8	84.8		2412	16	US-10-406-209-4	Sequence 4, Appli
43	17.8	84.8		2423	18	US-10-480-464-5	Sequence 5, Appli
44	17.8	84.8		2438	17	US-10-191-997-131	Sequence 131, App
45	17.8	84.8		2438	17	US-10-439-388-70	Sequence 70, Appl

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

ALIGNMENTS

```
RESULT 1
US-10-764-075-2

; Sequence 2, Application US/10764075
; Publication No. US20040265796A1
; GENERAL INFORMATION:
;  APPLICANT: Briese, Thomas
;  APPLICANT: Lipkin, Ian W.
;  APPLICANT: Palacios, Gustavo
;  APPLICANT: Jabado, Omar
;  TITLE OF INVENTION: Methods and Kits for Detecting SARS-Associated Coronavirus
;  FILE REFERENCE: 5199-87
;  CURRENT APPLICATION NUMBER: US/10/764,075
;  CURRENT FILING DATE: 2004-01-23
;  PRIOR APPLICATION NUMBER: Provisional Application 60/463,704
;  PRIOR FILING DATE: 2003-04-17
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn version 3.2
;  SEQ ID NO 2
;  LENGTH: 21
;  TYPE: DNA
;  ORGANISM: artificial sequence
;  FEATURE:
;  OTHER INFORMATION: primer
US-10-764-075-2

Query Match             100.0%;  Score 21;  DB 18;  Length 21;
Best Local Similarity   100.0%;  Pred. No. 1.1;
Matches   21;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 ATGACCACACAAGGCAGATGG 21
              |||||||||||||||||||||
Db          1 ATGACCACACAAGGCAGATGG 21
```

```
RESULT 2
US-10-764-075-1
; Sequence 1, Application US/10764075
; Publication No. US20040265796A1
; GENERAL INFORMATION:
;   APPLICANT: Briese, Thomas
;   APPLICANT: Lipkin, Ian W.
;   APPLICANT: Palacios, Gustavo
;   APPLICANT: Jabado, Omar
;   TITLE OF INVENTION: Methods and Kits for Detecting SARS-Associated Coronavirus
;   FILE REFERENCE: 5199-87
;   CURRENT APPLICATION NUMBER: US/10/764,075
;   CURRENT FILING DATE: 2004-01-23
;   PRIOR APPLICATION NUMBER: Provisional Application 60/463,704
;   PRIOR FILING DATE: 2003-04-17
;   NUMBER OF SEQ ID NOS: 16
;   SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
;   LENGTH: 1136
;   TYPE: DNA
;   ORGANISM: artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic nucleic acid sequence that includes the 3' non-coding
;   OTHER INFORMATION: region of the SARS associated coronavirus genome and a portion of
;   OTHER INFORMATION: the N gene of the SARS-associated coronavirus genome
US-10-764-075-1

  Query Match             100.0%;  Score 21;  DB 18;  Length 1136;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches    21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGACCACACAAGGCAGATGG 21
              |||||||||||||||||||||
Db        893 ATGACCACACAAGGCAGATGG 913


RESULT 3
US-10-699-936-16
; Sequence 16, Application US/10699936
; Publication No. US20050095582A1
; GENERAL INFORMATION:
;   APPLICANT: Gillim-Ross, Laura
;   APPLICANT: Taylor, Jill
;   APPLICANT: Scholl, David R.
;   APPLICANT: Wentworth, David E.
;   APPLICANT: Jollick, Joseph D.
;   TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory
;   TITLE OF INVENTION: Syndrome Coronavirus
;   FILE REFERENCE: DHI-07986
;   CURRENT APPLICATION NUMBER: US/10/699,936
;   CURRENT FILING DATE: 2003-11-03
;   NUMBER OF SEQ ID NOS: 87
;   SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16
;   LENGTH: 1620
;   TYPE: DNA
;   ORGANISM: SARS-CoV ZJ-HZ01
US-10-699-936-16

  Query Match             100.0%;  Score 21;  DB 19;  Length 1620;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches    21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGACCACACAAGGCAGATGG 21
              |||||||||||||||||||||
Db       1353 ATGACCACACAAGGCAGATGG 1373


RESULT 4
US-10-699-936-7
; Sequence 7, Application US/10699936
; Publication No. US20050095582A1
; GENERAL INFORMATION:
```

```
;   APPLICANT: Gillim-Ross, Laura
;   APPLICANT: Taylor, Jill
;   APPLICANT: Scholl, David R.
;   APPLICANT: Wentworth, David E.
;   APPLICANT: Jollick, Joseph D.
;   TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory
;   TITLE OF INVENTION: Syndrome Coronavirus
;   FILE REFERENCE: DHI-07986
;   CURRENT APPLICATION NUMBER: US/10/699,936
;   CURRENT FILING DATE: 2003-11-03
;   NUMBER OF SEQ ID NOS: 87
;   SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
;   LENGTH: 2304
;   TYPE: DNA
;   ORGANISM: SARS-CoV Shanghai LY
US-10-699-936-7

  Query Match             100.0%;  Score 21;  DB 19;  Length 2304;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches    21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGACCACACAAGGCAGATGG 21
              |||||||||||||||||||||
Db       2033 ATGACCACACAAGGCAGATGG 2053


RESULT 5
US-10-699-936-11
; Sequence 11, Application US/10699936
; Publication No. US2005009582A1
; GENERAL INFORMATION:
;   APPLICANT: Gillim-Ross, Laura
;   APPLICANT: Taylor, Jill
;   APPLICANT: Scholl, David R.
;   APPLICANT: Wentworth, David E.
;   APPLICANT: Jollick, Joseph D.
;   TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory
;   TITLE OF INVENTION: Syndrome Coronavirus
;   FILE REFERENCE: DHI-07986
;   CURRENT APPLICATION NUMBER: US/10/699,936
;   CURRENT FILING DATE: 2003-11-03
;   NUMBER OF SEQ ID NOS: 87
;   SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
;   LENGTH: 2304
;   TYPE: DNA
;   ORGANISM: SARS coronavirus Shanghai LY
US-10-699-936-11

  Query Match             100.0%;  Score 21;  DB 19;  Length 2304;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches    21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGACCACACAAGGCAGATGG 21
              |||||||||||||||||||||
Db       2033 ATGACCACACAAGGCAGATGG 2053


RESULT 6
US-10-889-447-3
; Sequence 3, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
;   APPLICANT: Bennett, C. Frank
;   APPLICANT: Jain, Ravi
;   TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
;   FILE REFERENCE: RTS-0685US
;   CURRENT APPLICATION NUMBER: US/10/889,447
;   CURRENT FILING DATE: 2004-07-12
;   PRIOR APPLICATION NUMBER: 60/486,670
;   PRIOR FILING DATE: 2003-07-12
;   NUMBER OF SEQ ID NOS: 241
```

; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 24774
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate BJ01
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-3

Query Match   100.0%;   Score 21; DB 19;   Length 24774;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGACCACACAAGGCAGATGG 21
          |||||||||||||||||||||
Db 24511 ATGACCACACAAGGCAGATGG 24531

RESULT 7
US-10-889-447-5
; Sequence 5, Application US/10889447
; Publication No. US2005007530 7A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 28920
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate BJ03
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-5

Query Match   100.0%;   Score 21; DB 19;   Length 28920;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGACCACACAAGGCAGATGG 21
          |||||||||||||||||||||
Db 28660 ATGACCACACAAGGCAGATGG 28680

RESULT 8
US-10-889-447-6
; Sequence 6, Application US/10889447
; Publication No. US2005007530 7A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 28920
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate BJ04

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-6

Query Match   100.0%;   Score 21; DB 19;   Length 28920;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGACCACACAAGGCAGATGG 21
          |||||||||||||||||||||
Db 28660 ATGACCACACAAGGCAGATGG 28680

RESULT 9
US-10-889-447-4
; Sequence 4, Application US/10889447
; Publication No. US2005007530 7A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 29291
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate BJ02
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-4

Query Match   100.0%;   Score 21; DB 19;   Length 29291;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGACCACACAAGGCAGATGG 21
          |||||||||||||||||||||
Db 29029 ATGACCACACAAGGCAGATGG 29049

RESULT 10
US-10-889-447-7
; Sequence 7, Application US/10889447
; Publication No. US2005007530 7A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 29430
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate GZ01
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-7

```
Query Match             100.0%;  Score 21;  DB 19;  Length 29430;
Best Local Similarity   100.0%;  Pred. No. 1.6;
Matches   21;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 ATGACCACCACAAGGCAGATGG 21
              |||||||||||||||||||||
Db      29163 ATGACCACCACAAGGCAGATGG 29183

RESULT 11
US-10-839-729-15
; Sequence 15, Application US/10839729
; Publication No. US20050002953A1
; GENERAL INFORMATION:
;  APPLICANT: Jens Herold
;  TITLE OF INVENTION: SARS-CORONAVIRUS VIRUS-LIKE PARTICLES
;  TITLE OF INVENTION:  AND METHODS OF USE
;  FILE REFERENCE: BIOBANK.013A
;  CURRENT APPLICATION NUMBER: US/10/839,729
;  CURRENT FILING DATE: 2004-05-04
;  PRIOR APPLICATION NUMBER: 60/468703
;  PRIOR FILING DATE: 2003-05-06
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
;  LENGTH: 29727
;  TYPE: DNA
;  ORGANISM: SARS Coronavirus
US-10-839-729-15

Query Match             100.0%;  Score 21;  DB 18;  Length 29727;
Best Local Similarity   100.0%;  Pred. No. 1.6;
Matches   21;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 ATGACCACCACAAGGCAGATGG 21
              |||||||||||||||||||||
Db      29398 ATGACCACCACAAGGCAGATGG 29418

RESULT 12
US-10-827-757-1
; Sequence 1, Application US/10827757
; Publication No. US20050004071A1
; GENERAL INFORMATION:
;  APPLICANT: Comper, Wayne
;  TITLE OF INVENTION:  Charged Polysaccharides Resistant To Lysosomal Degradation During
;  TITLE OF INVENTION:  Kidney Filtration And Renal Passage And Their Use To Treat Or
;  TITLE OF INVENTION:  Prevent Infection By Coronaviruses
;  FILE REFERENCE: 11213-007-999
;  CURRENT APPLICATION NUMBER: US/10/827,757
;  CURRENT FILING DATE: 2004-04-20
;  PRIOR APPLICATION NUMBER: 60/464,294
;  PRIOR FILING DATE: 2003-04-21
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
;  LENGTH: 29727
;  TYPE: DNA
;  ORGANISM: SARS-related coronavirus (Urbani strain)
US-10-827-757-1

Query Match             100.0%;  Score 21;  DB 18;  Length 29727;
Best Local Similarity   100.0%;  Pred. No. 1.6;
Matches   21;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 ATGACCACCACAAGGCAGATGG 21
              |||||||||||||||||||||
Db      29398 ATGACCACCACAAGGCAGATGG 29418

RESULT 13
US-10-889-447-8
```

```
; Sequence 8, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
;  APPLICANT: Bennett, C. Frank
;  APPLICANT: Jain, Ravi
;  TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
;  FILE REFERENCE: RTS-0685US
;  CURRENT APPLICATION NUMBER: US/10/889,447
;  CURRENT FILING DATE:  2004-07-12
;  PRIOR APPLICATION NUMBER: 60/486,670
;  PRIOR FILING DATE: 2003-07-12
;  NUMBER OF SEQ ID NOS: 241
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;  LENGTH: 29727
;  TYPE: DNA
;  ORGANISM: SARS coronavirus Urbani
US-10-889-447-8

Query Match             100.0%;  Score 21;  DB 19;  Length 29727;
Best Local Similarity   100.0%;  Pred. No. 1.6;
Matches   21;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 ATGACCACCACAAGGCAGATGG 21
              |||||||||||||||||||||
Db      29398 ATGACCACCACAAGGCAGATGG 29418

RESULT 14
US-10-699-936-1
; Sequence 1, Application US/10699936
; Publication No. US20050095582A1
; GENERAL INFORMATION:
;  APPLICANT: Gillim-Ross, Laura
;  APPLICANT:  Taylor, Jill
;  APPLICANT:  Scholl, David R.
;  APPLICANT:  Wentworth, David E.
;  APPLICANT:  Jollick, Joseph D.
;  TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory
;  TITLE OF INVENTION:  Syndrome Coronavirus
;  FILE REFERENCE: DHI-07986
;  CURRENT APPLICATION NUMBER: US/10/699,936
;  CURRENT FILING DATE: 2003-11-03
;  NUMBER OF SEQ ID NOS: 87
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
;  LENGTH: 29727
;  TYPE: DNA
;  ORGANISM: SARS coronavirus Urbani
US-10-699-936-1

Query Match             100.0%;  Score 21;  DB 19;  Length 29727;
Best Local Similarity   100.0%;  Pred. No. 1.6;
Matches   21;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 ATGACCACCACAAGGCAGATGG 21
              |||||||||||||||||||||
Db      29398 ATGACCACCACAAGGCAGATGG 29418

RESULT 15
US-10-839-729-17
; Sequence 17, Application US/10889729
; Publication No. US20050002953A1
; GENERAL INFORMATION:
;  APPLICANT: Jens Herold
;  TITLE OF INVENTION: SARS-CORONAVIRUS VIRUS-LIKE PARTICLES
;  TITLE OF INVENTION:  AND METHODS OF USE
;  FILE REFERENCE: BIOBANK.013A
;  CURRENT APPLICATION NUMBER: US/10/839,729
;  CURRENT FILING DATE: 2004-05-04
;  PRIOR APPLICATION NUMBER: 60/468703
;  PRIOR FILING DATE: 2003-05-06
```

; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 29736
; TYPE: DNA
; ORGANISM: SARS Coronavirus
US-10-839-729-17

Query Match   100.0%;   Score 21; DB 18; Length 29736;
Best Local Similarity   100.0%;   Pred. No. 1.6;
Matches   21;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 ATGACCACCACAAGGCAGATGG 21
           |||||||||||||||||||||||
Db   29383 ATGACCACCACAAGGCAGATGG 29403

Search completed: May 15, 2005, 16:48:44
Job time : 16.5004 secs

THIS PAGE LEFT BLANK

```
          GenCore version 5.1.6
          Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:   May 15, 2005, 11:03:36 ; Search time 92.4581 Seconds
                                    (without alignments)
                                    8645.536 Million cell updates/sec

Title:           US-10-764-075-2
Perfect score:   21
Sequence:        1 atgaccacacaaggcagatgg 21

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        34239544 seqs, 1903213470 residues

Total number of hits satisfying chosen parameters:   68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  EST:*
            1: gb_est1:*
            2: gb_est2:*
            3: gb_htc:*
            4: gb_est3:*
            5: gb_est4:*
            6: gb_est5:*
            7: gb_est6:*
            8: gb_gss1:*
            9: gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Match Length	DB	ID	Description
c	1	18.4	87.6	592	7	CO664875	CO664875 DG40-174i
c	2	18.4	87.6	1578	2	BF032477	BF032477 601452977
c	3	18	85.7	772	2	BF696011	BF696011 601852008
	4	18	85.7	1139	9	CL104513	CL104513 ISB1-43B2
	5	17.8	84.8	295	7	CN924279	CN924279 000414AEL
	6	17.8	84.8	329	7	T77190	T77190 yc95b09.r1
	7	17.8	84.8	356	1	AL119881	AL119881 DKFZp761M
	8	17.8	84.8	456	8	AQ669320	AQ669320 HS_5364_A
	9	17.8	84.8	482	8	AQ133467	AQ133467 HS_3039_A
c	10	17.8	84.8	486	8	AQ880134	AQ880134 HS_5037_B
	11	17.8	84.8	506	7	R17506	R17506 yg02e12.r1
c	12	17.8	84.8	536	5	BI536734	BI536734 396079 MA
	13	17.8	84.8	543	9	CE836567	CE836567 tigr-gss-
	14	17.8	84.8	547	2	CF804618	CF804618 1ad76d11.
c	15	17.8	84.8	558	4	BJ200351	BJ200351 BJ200351
	16	17.8	84.8	606	7	CN853628	CN853628 Ha_mx0_15
	17	17.8	84.8	651	9	CE295750	CE295750 tigr-gss-
	18	17.8	84.8	659	4	BM015637	BM015637 603641872
	19	17.8	84.8	689	9	CE457443	CE457443 tigr-gss-
	20	17.8	84.8	721	4	BG768359	BG768359 60274401S
	21	17.8	84.8	772	2	BE281640	BE281640 601155157
	22	17.8	84.8	886	5	BQ678947	BQ678947 AGENCOURT
	23	17.8	84.8	894	5	BX389435	BX389435 BX389435
	24	17.8	84.8	919	5	BQ953874	BQ953874 AGENCOURT
	25	17.8	84.8	919	5	BU179230	BU179230 AGENCOURT
	26	17.8	84.8	957	2	BF347731	BF347731 602022564
	27	17.8	84.8	970	4	BG251251	BG251251 602365877
	28	17.8	84.8	982	5	BQ680344	BQ680344 AGENCOURT
	29	17.8	84.8	1017	5	BU167886	BU167886 AGENCOURT
c	30	17.8	84.8	1125	5	BX440072	BX440072 BX440072
	31	17.8	84.8	1157	5	BQ955716	BQ955716 AGENCOURT
	32	17.4	82.9	2438	2	BC045610	BC045610 Homo sapi
c	33	17.4	82.9	410	2	BE627254	BE627254 ut86a09.x
c	34	17.4	82.9	440	8	AQ221669	AQ221669 HS_2016_A
	35	17.4	82.9	445	1	AI426376	AI426376 mm48c01.x
c	36	17.4	82.9	445	8	AQ244968	AQ244968 HS_3204_B
	37	17.4	82.9	560	4	BM822423	BM822423 K-EST0092
c	38	17.4	82.9	695	9	CE631183	CE631183 tigr-gss-
	39	17.4	82.9	705	8	CC004155	CC004155 PUDEP40TD
	40	17.4	82.9	847	8	CC392087	CC392087 PUHSM35TB
	41	17.4	82.9	3606	2	AK036023	AK036023 Mus muscu
	42	17	81.0	371	4	BG009169	BG009169 ILS-GN024
c	43	17	81.0	457	8	AQ666653	AQ666653 HS_2111_A
c	44	17	81.0	529	2	BE010668	BE010668 MR0-BN018
c	45	17	81.0	576	2	AV975852	AV975852 AV975852

ALIGNMENTS

RESULT 1
CO664875/c

```
LOCUS       CO664875         592 bp    mRNA    linear   EST 23-JUL-2004
DEFINITION  DG40-174i12 DG40-uterus Canis familiaris cDNA 3', mRNA sequence.
ACCESSION   CO664875
VERSION     CO664875.1  GI:50581665
KEYWORDS    EST.
SOURCE      Canis familiaris (dog)
ORGANISM    Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1 (bases 1 to 592)
  AUTHORS   Schlueter,T., Hermanns,J., Weindel,M., Schuette,D., Kranz,H.,
            Henrich,J. and Loebbert,R.
  TITLE     Dog arrayTAG cDNA clone collection
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Thomas Schlueter
            LION bioscience AG
            Waldhoferstrasse 98, D-69123 Heidelberg, Germany
            Tel: +49 6221 4038 150
            Fax: +49 6221 4038 290
            Email: Thomas.Schlueter@lionbioscience.com.
FEATURES             Location/Qualifiers
     source          1..592
                     /organism="Canis familiaris"
                     /mol_type="mRNA"
                     /strain="Beagle"
                     /db_xref="taxon:9615"
                     /tissue_type="uterus"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="DG40-uterus"
                     /note="organ: uterus; Vector: Dog pBluescript LION"
ORIGIN

Query Match             87.6%;  Score 18.4;  DB 7;  Length 592;
  Best Local Similarity   95.0%;  Pred. No. 5.6e+02;
  Matches   19;  Conservative    0;  Mismatches   1;  Indels    0;  Gaps    0;

Qy      2 TGACCACACAAGGCAGATGG 21
          |||||||||||||| |||||
Db    525 TGACCACACAAGGGAGATGG 506
```

RESULT 2
BF032477/c

```
LOCUS       BF032477        1578 bp    mRNA    linear   EST 20-OCT-2000
```

```
DEFINITION   60145297F1 NIH_MGC_66 Homo sapiens cDNA clone IMAGE:3856626 5',
             mRNA sequence.
ACCESSION    BF032477
VERSION      BF032477.1  GI:10740177
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 1578)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLAM9585 row: j column: 19
             High quality sequence start: 23
             High quality sequence stop: 182.
FEATURES             Location/Qualifiers
     source          1..1578
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3856626"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_66"
                     /note="Organ: ovary; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.8 kb. Library constructed by Life
                     Technologies. "
ORIGIN

Query Match              87.6%;  Score 18.4;  DB 2;  Length 1578;
  Best Local Similarity  90.5%;  Pred. No. 6.3e+02;
  Matches  19;  Conservative   0;  Mismatches  2;  Indels   0;  Gaps   0;

Qy          1 ATGACCACACAAGGCAGATGG 21
              ||||||| ||||||| ||||
Db        652 ATGACCACNCAAGGCATATGG 632

RESULT 3
BF696011/c
LOCUS       BF696011                 772 bp    mRNA    linear   EST 22-DEC-2000
DEFINITION   601852008F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4076017 5',
             mRNA sequence.
ACCESSION    BF696011
VERSION      BF696011.1  GI:11981430
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 772)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: ATCC
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:

             http://image.llnl.gov
             Plate: LLCM928 row: p column: 02
             High quality sequence stop: 765.
FEATURES             Location/Qualifiers
     source          1..772
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4076017"
                     /tissue_type="primitive neuroectoderm"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_56"
                     /note="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggcgcctcggcc); Site 2: SfiI (ggccattatggcc);
                     Double-stranded cDNA was prepared from cell line RNA. 5'
                     and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence: 5'-ATTCTAGAGGCGGCCGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C, G, or T). Average
                     insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
                     contained inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA)."
ORIGIN

Query Match              85.7%;  Score 18;  DB 2;  Length 772;
  Best Local Similarity 100.0%;  Pred. No. 9e+02;
  Matches  18;  Conservative   0;  Mismatches  0;  Indels   0;  Gaps   0;

Qy          4 ACCACACAAGGCAGATGG 21
              ||||||||||||||||||
Db        590 ACCACACAAGGCAGATGG 573

RESULT 4
CL104513
LOCUS       CL104513                1139 bp    DNA     linear   GSS 05-JAN-2004
DEFINITION   ISB1-43B21_T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-43B21,
             genomic survey sequence.
ACCESSION    CL104513
VERSION      CL104513.1  GI:40599148
KEYWORDS     GSS.
SOURCE       Xenopus tropicalis (western clawed frog)
ORGANISM     Xenopus tropicalis
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
             Xenopodinae; Xenopus; Silurana.
REFERENCE    1 (bases 1 to 1139)
AUTHORS      Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
             Mardis,E. and Wilson,R.
TITLE        A physical map of the xenopus tropicalis genome
JOURNAL      Unpublished (2003)
COMMENT      Contact: Richard K Wilson
             Genome Sequencing Center
             Washington University School of Medicine
             Email: submissions@watson.wustl.edu
             Insert Length: 75000   Std Error: 0.00
             Seq primer: T7 TAATACGACTCACTATAGGG
             Class: BAC ends
             High quality sequence start: 25
             High quality sequence stop: 765.
FEATURES             Location/Qualifiers
     source          1..1139
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:8364"
                     /clone="ISB1-43B21"
                     /clone_lib="ISB1"
                     /note="Vector: pBeloBAC11; ISB-1 Xenopus tropicalis BAC
                     Library Segment 1"
ORIGIN

Query Match              85.7%;  Score 18;  DB 9;  Length 1139;
```

Best Local Similarity   100.0%;   Pred. No. 9.5e+02;
Matches   18;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 ATGACCACACAAGGCAGA 18
            ||||||||||||||||||
Db   323 ATGACCACACAAGGCAGA 340

RESULT 5
CN924279
LOCUS   CN924279   295 bp   mRNA   linear   EST 07-JUN-2004
DEFINITION   000414AELA009380HT (AELA) Royal Gala young expanding leaf Malus x
             domestica cDNA clone AELA009380, mRNA sequence.
ACCESSION   CN924279
VERSION   CN924279.1 GI:48397092
KEYWORDS   EST.
SOURCE   Malus x domestica (cultivated apple)
ORGANISM   Malus x domestica
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
             rosids; eurosids I; Rosales; Rosaceae; Maloideae; Malus.
REFERENCE   1 (bases 1 to 295)
AUTHORS   Beuning,L., Bowen,J., Crowhurst,R., Gleave,A., Janssen,B.,
             McArtney,S., Newcomb,R., Ross,G., Snowden,K., Walton,E. and Yauk,Y.
TITLE   HortResearch Apple EST Project
JOURNAL   Unpublished (2004)
COMMENT   Contact: Gleave,A.
             The Horticulture and Food Research Institute of New Zealand Ltd
             Sequencing Facility
             120 Mt Albert Rd, Mt Albert, Auckland, New Zealand
             Tel: 00 64 09 815 4200
             Fax: 00 64 09 815 4201
             Email: est@hortresearch.co.nz.
FEATURES   Location/Qualifiers
     source   1..295
                     /organism="Malus x domestica"
                     /mol_type="mRNA"
                     /db_xref="taxon:3750"
                     /clone="AELA009380"
                     /tissue_type="Leaf"
                     /dev_stage="Young, expanding"
                     /clone_lib="(AELA) Royal Gala young expanding leaf"
                     /note="Vector: pBK-CMV; Library sequenced by Genesis
                     Research & Development"
ORIGIN

Query Match   84.8%;   Score 17.8; DB 7; Length 295;
Best Local Similarity   90.5%;   Pred. No. 1e+03;
Matches   19;   Conservative   0;   Mismatches   2;   Indels   0;   Gaps   0;

Qy   1 ATGACCACACAAGGCAGATGG 21
            ||||||| |||||||||||||
Db   49 ATGACCATAGAAGGCAGATGG 69

RESULT 6
T77190
LOCUS   T77190   329 bp   mRNA   linear   EST 06-MAR-1995
DEFINITION   yc95b09.r1 Soares infant brain 1NIB Homo sapiens cDNA clone
             IMAGE:23715 5' similar to SP:COG2_MOUSE P33434 72 KD TYPE IV
             COLLAGENASE PRECURSOR ;, mRNA sequence.
ACCESSION   T77190
VERSION   T77190.1 GI:694393
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 329)
AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
             Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
             Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
             Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and

---

Wilson,R.
TITLE   The WashU-Merck EST Project
JOURNAL   Unpublished (1995)
COMMENT   Contact: Wilson RK
             Washington University School of Medicine
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: est@watson.wustl.edu
             Insert Size: 1466
             High quality sequence stops: 249 Source: IMAGE Consortium, LLNL
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             Insert Length: 1466   Std Error: 0.00
             Seq primer: M13RP1
             High quality sequence stop: 249.
FEATURES   Location/Qualifiers
     source   1..329
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:396062"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:23715"
                     /sex="female"
                     /dev_stage="73 days post natal"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares infant brain 1NIB"
                     /note="Organ: whole brain; Vector: Lafmid BA; Site_1: Not
                     I; Site_2: Hind III; 1st strand cDNA was primed with a Not
                     I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTT 3'];
                     double-stranded cDNA was ligated to Hind III adaptors
                     (Pharmacia), digested with Not I and directionally cloned
                     into the Not I and Hind III sites of the Lafmid BA vector.
                     Library went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
ORIGIN

Query Match   84.8%;   Score 17.8; DB 7; Length 329;
Best Local Similarity   90.5%;   Pred. No. 1e+03;
Matches   19;   Conservative   0;   Mismatches   2;   Indels   0;   Gaps   0;

Qy   1 ATGACCACACAAGGCAGATGG 21
            |||||||||||||| ||| |||
Db   212 ATGACCACAGGAGGCACATGG 232

RESULT 7
AL119881
LOCUS   AL119881   356 bp   mRNA   linear   EST 04-SEP-2003
DEFINITION   DKFZp761M2224_r1 761 (synonym: hamy2) Homo sapiens cDNA clone
             DKFZp761M2224 5', mRNA sequence.
ACCESSION   AL119881
VERSION   AL119881.1 GI:5925780
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 356)
AUTHORS   Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and
             Wiemann,S.
TITLE   EST (Bloecker, et al.)
JOURNAL   Unpublished (1999)
COMMENT   Contact: MIPS
             MIPS
             Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
             This is the 5' sequence of the clone insert
             Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
             Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
             sequenced by GBF (National Research Centre for Biotechnology Ltd.,
             Braunschweig/Germany) within the cDNA sequencing consortium of the
             German Genome Project.

No s1 sequence available.
This clone (DKFZp761M2224) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES   Location/Qualifiers
     source   1..356
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="DKFZp761M2224"
                  /tissue_type="amygdala"
                  /dev_stage="adult"
                  /lab_host="DH10B"
                  /clone_lib="761 (synonym: hamy2)"
                  /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match   84.8%;   Score 17.8; DB 1; Length 356;
Best Local Similarity   90.5%; Pred. No. 1.1e+03;
Matches 19;   Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 ATGACCACCACAAGGCAGATGG 21
          |||||||||| |||| ||||
Db   210 ATGACCACCACGAGGCACATGG 230

RESULT 8
AQ669320
LOCUS   AQ669320   456 bp   DNA   linear   GSS 24-JUN-1999
DEFINITION   HS_5364_A1_E02_T7A RPCI-11 Human Male BAC Library Homo sapiens
             genomic clone Plate=940 Col=3 Row=I, genomic survey sequence.
ACCESSION   AQ669320
VERSION   AQ669320.1   GI:5202066
KEYWORDS   GSS.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 456)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
             Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
             Hood,L.
  TITLE   Sequence-tagged connectors: A sequence approach to mapping and
             scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
   PUBMED   10449764
COMMENT   Contact: Mahairas GG, Wallace JC, Hood L
             High Throughput Sequencing Center
             University of Washington
             401 Queen Anne Avenue North, Seattle, WA 98109, USA
             Tel: (206) 616-3618
             Fax: (206) 616-3887
             Email: jwallace@u.washington.edu
             Clones are derived from the human BAC library RPCI-11. For BAC
             library availability, please contact Pieter de Jong
             (pieter@dejong.med.buffalo.edu). Clones may be purchased from
             BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
             or from Resear h Genetics (info@resgen.com). BAC end Web Server:
             http://www.htsc.washington.edu
             Plate: 940 row: I column: 3
             Seq primer: T7
             Class: BAC ends
             High quality sequence stop: 456.
FEATURES   Location/Qualifiers
     source   1..456
                  /organism="Homo sapiens"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:9606"
                  /clone="Plate=940 Col=3 Row=I"
                  /sex="male"
                  /clone_lib="RPCI-11 Human Male BAC Library"
                  /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;

Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACe3.6 vector at EcoRI sites"

ORIGIN

Query Match   84.8%;   Score 17.8; DB 8; Length 456;
Best Local Similarity   90.5%; Pred. No. 1.1e+03;
Matches 19;   Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 ATGACCACCACAAGGCAGATGG 21
          |||||||||| |||| ||||
Db   357 ATGACTACTCAAGGCAGATGG 377

RESULT 9
AQ133467
LOCUS   AQ133467   482 bp   DNA   linear   GSS 23-SEP-1998
DEFINITION   HS_3039_A2_E05 MR CIT Approved Human Genomic Sperm Library D Homo
             sapiens genomic clone Plate=3039 Col=10 Row=I, genomic survey
             sequence.
ACCESSION   AQ133467
VERSION   AQ133467.1   GI:3524833
KEYWORDS   GSS.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 482)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
             Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
             Hood,L.
  TITLE   Sequence-tagged connectors: A sequence approach to mapping and
             scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
   PUBMED   10449764
COMMENT   Contact: Mahairas GG, Wallace JC, Hood L
             High Throughput Sequencing Center
             University of Washington
             401 Queen Anne Avenue North, Seattle, WA 98109, USA
             Tel: (206) 616-3618
             Fax: (206) 616-3887
             Email: jwallace@u.washington.edu
             Sequence Tagged Connector
             Plate: 3039 row: I column: 10
             Class: BAC ends
             High quality sequence stop: 482.
FEATURES   Location/Qualifiers
     source   1..482
                  /organism="Homo sapiens"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:9606"
                  /clone="Plate=3039 Col=10 Row=I"
                  /sex="male"
                  /clone_lib="CIT Approved Human Genomic Sperm Library D"
                  /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                  E-Coli DH10B"

ORIGIN

Query Match   84.8%;   Score 17.8; DB 8; Length 482;
Best Local Similarity   90.5%; Pred. No. 1.1e+03;
Matches 19;   Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 ATGACCACCACAAGGCAGATGG 21
          |||||||||| |||| ||||
Db   161 ATGACTACTCAAGGCAGATGG 181

RESULT 10
AQ880134/c
LOCUS   AQ880134   486 bp   DNA   linear   GSS 09-NOV-1999
DEFINITION   HS_5037_B1_A04_T7 RPCI-11 Human Male BAC Library Homo sapiens

ACCESSION   AQ880134
VERSION   AQ880134.1   GI:6311601
KEYWORDS   GSS.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 466)
AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
             Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
             Hood,L.
TITLE   Sequence-tagged connectors: A sequence approach to mapping and
             scanning the human genome
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE   99380589
PUBMED   10449764
COMMENT   Contact: Mahairas GG, Wallace JC, Hood L
             High Throughput Sequencing Center
             University of Washington
             401 Queen Anne Avenue North, Seattle, WA 98109, USA
             Tel: (206) 616-3618
             Fax: (206) 616-3887
             Email: jwallace@u.washington.edu
             Clones are derived from the human BAC library RPCI-11. For BAC
             library availability, please contact Pieter de Jong
             (pieter@dejong.med.buffalo.edu). Clones may be purchased from
             BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
             or from Resear h Genetics (info@resgen.com). BAC end Web Server:
             http://www.htsc.washington.edu
             Plate: 8805 row: B column: 7
             Seq primer: T7
             Class: BAC ends
             High quality sequence stop: 486.
FEATURES   Location/Qualifiers
     source   1..486
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=8805 Col=7 Row=B"
                     /sex="male"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
ORIGIN

Query Match   84.8%; Score 17.8; DB 8; Length 486;
Best Local Similarity   90.5%; Pred. No. 1.1e+03;
Matches   19; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 ATGACCACCAAGGCAGATGG 21
              |||||||| |||||||||||
Db   411 ATGACCATTCAAGGCAGATGG 391

RESULT 11
R17506
LOCUS   R17506   506 bp   mRNA   linear   EST 14-APR-1995
DEFINITION   yg02e12.r1 Soares infant brain 1NIB Homo sapiens cDNA clone
             IMAGE:30850 5', mRNA sequence.
ACCESSION   R17506
VERSION   R17506.1   GI:771116
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 506)
AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
             Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,

---

Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
             Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
             Wilson,R.
TITLE   The WashU-Merck EST Project
JOURNAL   Unpublished (1995)
COMMENT   Contact: Wilson RK
             Washington University School of Medicine
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: est@watson.wustl.edu
             Insert Size: 1446
             High quality sequence stops: 332
             Source: IMAGE Consortium, LLNL
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             Insert Length: 1446   Std Error: 0.00
             Seq primer: M13RP1
             High quality sequence stop: 332.
FEATURES   Location/Qualifiers
     source   1..506
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:403197"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30850"
                     /sex="female"
                     /dev_stage="73 days post natal"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares infant brain 1NIB"
                     /note="Organ: whole brain; Vector: Lafmid BA; Site 1: Not
                     I; Site_2: Hind III; 1st strand cDNA was primed with a Not
                     I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'];
                     double-stranded cDNA was ligated to Hind III adaptors
                     (Pharmacia), digested with Not I and directionally cloned
                     into the Not I and Hind III sites of the Lafmid BA vector.
                     Library went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
ORIGIN

Query Match   84.8%; Score 17.8; DB 7; Length 506;
Best Local Similarity   90.5%; Pred. No. 1.1e+03;
Matches   19; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 ATGACCACCAAGGCAGATGG 21
              ||||||||| |||||| |||
Db   291 ATGACCACCACGGGCACATGG 311

RESULT 12
BI536734/c
LOCUS   BI536734   536 bp   mRNA   linear   EST 30-AUG-2001
DEFINITION   396079 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   BI536734
VERSION   BI536734.1   GI:15377844
KEYWORDS   EST.
SOURCE   Bos taurus (cow)
ORGANISM   Bos taurus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
             Bovinae; Bos.
REFERENCE   1 (bases 1 to 536)
AUTHORS   Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
             Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C.,
             Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A.,
             Chitko-McKown,C.G., Pertea,G., Holt,I., Karamycheva,S., Liang,F.,
             Quackenbush,J. and Keele,J.W.
TITLE   Sequence evaluation of four pooled-tissue normalized bovine cDNA
             libraries and construction of a gene index for cattle
JOURNAL   Genome Res. 11 (4), 626-630 (2001)
MEDLINE   21180013
PUBMED   11282978

COMMENT   Contact: Smith TPL
              USDA, ARS, US Meat Animal Research Center
              PO Box 166, Clay Center, NE 68933-0166, USA
              Tel: 402 762 4366
              Fax: 402 762 4390
              Email: smith@email.marc.usda.gov
              Single pass sequencing. Bases called and alt trimmed with phred
              v0.980904.e. Vector identified by cross_match with the -minscore 18
              and -minmatch 12 options.
              PCR PRimers
              FORWARD: AGGAAACAGCTATGACCAT
              BACKWARD: GTTTTCCCAGTCACGACG
              Plate: 119 row: O column: 2
              Seq primer: ATTTAGGTGACACTATAG.
FEATURES   Location/Qualifiers
     source   1..536
              /organism="Bos taurus"
              /mol_type="mRNA"
              /db_xref="taxon:9913"
              /tissue_type="pooled"
              /lab_host="DH10B"
              /clone_lib="MARC 4BOV"
              /note="vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
              Library made from pooled tissue from day 20 and day 40
              embryos."
ORIGIN

Query Match   84.8%; Score 17.8; DB 4; Length 536;
Best Local Similarity   90.5%; Pred. No. 1.1e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy   1 ATGACCACACAAGGCAGATGG 21
             ||||||| |||||||||||||
Db   28 ATGACCACGCAAGGCAGATAG 8

RESULT 13
CE836567
CE836567
LOCUS   CE836567   543 bp   DNA   linear   GSS 30-SEP-2003
DEFINITION tigr-gss-dog-17000332589366 Dog Library Canis familiaris genomic,
            genomic survey sequence.
ACCESSION   CE836567
VERSION   CE836567.1 GI:37177587
KEYWORDS   GSS.
SOURCE   Canis familiaris (dog)
  ORGANISM Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1 (bases 1 to 543)
  AUTHORS   Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
            Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
            Venter,J.C.
  TITLE   The dog genome: survey sequencing and comparative analysis
  JOURNAL   Science 301 (5641), 1898-1903 (2003)
   MEDLINE 22875432
    PUBMED 14512627
COMMENT   Contact: Kirkness EF
            The Institute for Genomic Research
            Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
            Rockville, MD 20850, USA
            Tel: 301-838-0200
            Fax: 301-838-0208
            Email: ekirkness@tigr.org
            Class: shotgun.
FEATURES   Location/Qualifiers
     source   1..543
              /organism="Canis familiaris"
              /mol_type="genomic DNA"
              /strain="Standard Poodle"
              /db_xref="taxon:9615"
              /clone_lib="Dog Library"
              /note="Site 1: BstXI; Libraries were prepared from
              peripheral blood"

ORIGIN

Query Match   84.8%; Score 17.8; DB 9; Length 543;
Best Local Similarity   90.5%; Pred. No. 1.1e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy   1 ATGACCACACAAGGCAGATGG 21
             ||| |||| |||||||||||||
Db   367 ATGTCCACATAAGGCAGATGG 387

RESULT 14
CF804618
CF804618
LOCUS   CF804618   547 bp   mRNA   linear   EST 24-OCT-2003
DEFINITION 1ad76d11.y1 Gastric Epithelial Progenitor 2 Mus musculus cDNA 5'
            similar to SW:ROA1_BOVIN P09867 HETEROGENEOUS NUCLEAR
            RIBONUCLEOPROTEIN A1 ;, mRNA sequence.
ACCESSION   CF804618
VERSION   CF804618.1 GI:37973438
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 547)
  AUTHORS   Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J.,
            Wylie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennet,J.,
            Ronko,I., Tsagareishvili,R., Belaygorod,L., Grow,A., Maguire,L.,
            Waterston,R. and Wilson,R.
  TITLE   WashU Stem cell EST Project
  JOURNAL   Unpublished (2002)
COMMENT   Contact: Jeff Gordon and Mike Lovett
            WashU, Human Genetics Division
            Washington University School of Medicine
            1st strand of cDNA was synthesized with reverse transcriptase and
            oligo(dT) beads, then cDNA was amplified by PCR using modified
            SMART primers. The final cDNA was cloned in pAMP1 vector in
            annealing reaction with Uracil DNA Glycosylase (UDG). Library
            constructed by Y.Korshunova and M. Lovett. Library materials
            provided by Mills JC & Gordon JI.
            Putative full length read
            vector to vector length is
            Seq primer: -40RP from Gibco.
FEATURES   Location/Qualifiers
     source   1..547
              /organism="Mus musculus"
              /mol_type="mRNA"
              /db_xref="taxon:10090"
              /tissue_type="Laser-captured isthmal cells from tox176
              transgenic mice"
              /dev_stage="adult"
              /lab_host="DH5alpha"
              /clone_lib="Gastric Epithelial Progenitor 2"
              /note="Vector: pAMP1; 1st strand of cDNA was synthesized
              with reverse transcriptase and oligo(dT) beads, then cDNA
              was amplified by PCR using modified SMART primers. The
              final cDNA was cloned in pAMP1 vector in annealing
              reaction with Uracil DNA Glycosylase (UDG). Library
              constructed by Y.Korshunova and M. Lovett. Library
              materials provided by Mills JC & Gordon JI."

ORIGIN

Query Match   84.8%; Score 17.8; DB 7; Length 547;
Best Local Similarity   90.5%; Pred. No. 1.1e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy   1 ATGACCACACAAGGCAGATGG 21
             | |||||||||| ||||||
Db   372 AAGACCACACAAGGTAGATGG 392

RESULT 15
BJ200351/c

```
LOCUS       BJ200351              558 bp    mRNA    linear   EST 21-OCT-2003
DEFINITION  BJ200351 normalized full length cDNA library, chloronemata,
            caulonemata and rhizoid-like protonemata Physcomitrella patens
            subsp. patens cDNA clone pphn3g23 5', mRNA sequence.
ACCESSION   BJ200351
VERSION     BJ200351.1  GI:18368266
KEYWORDS    EST.
SOURCE      Physcomitrella patens subsp. patens
  ORGANISM  Physcomitrella patens subsp. patens
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
            Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
REFERENCE   1  (bases 1 to 558)
  AUTHORS   Nishiyama,T., Fujita,T., Shin-i,T., Seki,M., Nishide,H.,
            Uchiyama,I., Kamiya,A., Carninci,P., Hayashizaki,Y., Shinozaki,K.,
            Kohara,Y. and Hasebe,M.
  TITLE     Comparative genomics of Physcomitrella patens gametophytic
            transcriptome and Arabidopsis thaliana: implication for land plant
            evolution
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
  MEDLINE   22709184
   PUBMED   12808149
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp
            A backbone of the vector is basically from pBluescript II (KS),
            that was in vivo excised from a l-FLC phage vector (Carninci et al.
            2001). 5' end of the cDNA that was digested with XhoI was ligated
            to SalI site of the vector and the 3' end including polyA tail was
            ligated to BamHI site of the
            vector(5'- gAgAgAgAggATCCAACCCTggAgAgTTTTTTTTTTTTTTVN-3' was
            used as a 1st 3' primer, and
            5'-ggTTCTCgAgTCATCgCTgTTCCAgACAgCgATgACTCgAACCgNNNNN-3' as 2nd
            5'-hairpin primer, giving the following 5' boarder sequence,
            AGGCCAAATCGGCCGAGCTCGAATTCGTCGAGAACCG). cDNA instert could be
            amplified with conventional T7 and T3 primers. This full-length
            cDNA library was generated according to the method described in
            Nishiyama et al. (2003).
            Protonemata were blended by the POLYTRON, and then cultivated on
            the BCDATG medium for 13- 14 days under the continuous light.
            These clones are available from RIKEN Bio Resource Center
            (http://www.brc.riken.go.jp/lab/epd/Eng/index.html). The database
            of Physcomitrella EST clones is available at the PHYSCObase
            (http://moss.nibb.ac.jp).
FEATURES             Location/Qualifiers
     source          1..558
                     /organism="Physcomitrella patens subsp. patens"
                     /mol_type="mRNA"
                     /sub_species="patens"
                     /db_xref="taxon:145481"
                     /clone="pphn3g23"
                     /tissue_type="mixture of chloronemata, caulonemata and
                     rhizoid-like protonemata"
                     /clone_lib="normalized full length cDNA library,
                     chloronemata, caulonemata and rhizoid-like protonemata"
ORIGIN
Query Match         84.8%;  Score 17.8;  DB 4;  Length 558;
Best Local Similarity  90.5%;  Pred. No. 1.1e+03;
Matches  19;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy     1 ATGACCACACAAGGCAGATCG 21
           ||||| ||||||| |||||||
Db    49 ATGACAACACAAAGCAGATGG 29
```

Search completed: May 15, 2005, 14:18:45
Job time : 100.458 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 11:00:45 ; Search time 5188.09 Seconds
                                            (without alignments)
                                            10609.895 Million cell updates/sec

Title: US-10-764-075-1
Perfect score: 1136
Sequence: 1 aggcatcgtatgggttgcaa.........tgaataatgctagggagagc 1136

Scoring table: IDENTITY NUC
               Gapop 10_0 , Gapext 1.0

Searched: 4708233 seqs, 2422760795 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

                 SUMMARIES

Database :   GenEmbl:*
         1: gb_ba:*
         2: gb_htg:*
         3: gb_in:*
         4: gb_om:*
         5: gb_ov:*
         6: gb_pat:*
         7: gb_ph:*
         8: gb_pl:*
         9: gb_pr:*
        10: gb_ro:*
        11: gb_sts:*
        12: gb_sy:*
        13: gb_un:*
        14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description
1	1136	100.0	1873	14	AY534758S4	AY534761 SARS coro
2	1136	100.0	1938	14	AY534762S6	AY534767 SARS coro
3	1136	100.0	2810	14	AY290752	AY290752 SARS coro
4	1136	100.0	8581	14	AY304487	AY304487 SARS coro
5	1136	100.0	11010	14	AY304493	AY304493 SARS coro
6	1136	100.0	11010	14	AY304494	AY304494 SARS coro
7	1136	100.0	13471	14	AY304492	AY304492 SARS coro
8	1136	100.0	29013	14	AY463060	AY463060 SARS coro
9	1136	100.0	29350	14	AY394999	AY394999 SARS coro
10	1136	100.0	29350	14	AY395000	AY395000 SARS coro
11	1136	100.0	29350	14	AY395001	AY395001 SARS coro
12	1136	100.0	29350	14	AY395002	AY395002 SARS coro
13	1136	100.0	29530	14	AY394985	AY394985 SARS coro
14	1136	100.0	29573	14	AY338174	AY338174 SARS coro
15	1136	100.0	29573	14	AY338175	AY338175 SARS coro
16	1136	100.0	29573	14	AY348314	AY348314 SARS coro
17	1136	100.0	29577	14	AY559094	AY559094 SARS coro
18	1136	100.0	29592	14	AY463059	AY463059 SARS coro
19	1136	100.0	29620	14	AY395004	AY395004 SARS coro
20	1136	100.0	29640	14	AY394978	AY394978 SARS coro
21	1136	100.0	29645	14	AY394979	AY394979 SARS coro
22	1136	100.0	29646	14	AY394982	AY394982 SARS coro
23	1136	100.0	29661	14	AY559086	AY559086 SARS coro
24	1136	100.0	29670	14	AY559082	AY559082 SARS coro
25	1136	100.0	29699	14	AY394983	AY394983 SARS coro
26	1136	100.0	29705	14	AY283795	AY283795 SARS coro
27	1136	100.0	29706	14	AY283797	AY283797 SARS coro
28	1136	100.0	29709	14	AY394987	AY394987 SARS coro
29	1136	100.0	29710	14	AY559091	AY559091 SARS coro
30	1136	100.0	29711	14	AY283794	AY283794 SARS coro
31	1136	100.0	29711	14	AY283796	AY283796 SARS coro
32	1136	100.0	29711	14	AY283798	AY283798 SARS coro
33	1136	100.0	29711	14	AY427439	AY427439 SARS coro
34	1136	100.0	29712	14	AY559093	AY559093 SARS coro
35	1136	100.0	29713	14	AY559085	AY559085 SARS coro
36	1136	100.0	29713	14	AY559092	AY559092 SARS coro
37	1136	100.0	29714	14	AY559088	AY559088 SARS coro
38	1136	100.0	29715	14	AY461660	AY461660 SARS coro
39	1136	100.0	29715	14	AY559097	AY559097 SARS coro
40	1136	100.0	29716	14	AY559081	AY559081 SARS coro
41	1136	100.0	29716	14	AY559087	AY559087 SARS coro
42	1136	100.0	29716	14	AY595412	AY595412 SARS coro
43	1136	100.0	29719	14	AY559095	AY559095 SARS coro
44	1136	100.0	29720	14	AY304495	AY304495 SARS coro
45	1136	100.0	29720	14	AY559096	AY559096 SARS coro

ALIGNMENTS

RESULT 1
AY534758S4
LOCUS   AY534758S4   1873 bp   RNA   linear   VRL 17-MAR-2004
DEFINITION SARS coronavirus Sin0409, partial sequence.
ACCESSION   AY534761
VERSION   AY534761.1 GI:45384968
KEYWORDS
SEGMENT   4 of 4
SOURCE   SARS coronavirus Sin0409
ORGANISM   SARS coronavirus Sin0409
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1 (bases 1 to 1873)
AUTHORS   Lim,P.L., Kurup,A., Gopalakrishna,G., Chan,K.P., Wong,C.W.,
            Ng,L.C., Se-Thoe,S.Y., Oon,L., Bai,X., Stanton,L.W., Ruan,Y.,
            Miller,L.D., Vega,V.B., James,L., Ooi,P.L., Kai,C.S., Olsen,S.J.,
            Ang,B. and Leo,Y.S.
TITLE   Laboratory-acquired severe acute respiratory syndrome (SARS) -
            Singapore 2003
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 1873)
AUTHORS   Wei,C.L., Lee,C., Lin,S., Thoreau,H., Vega,V.B., Stanton,L.W. and
            Ruan,Y.
TITLE   Direct Submission
JOURNAL   Submitted (28-JAN-2004) Genome Institute of Singapore, 60, Biopolis
            Street, 02-01, Genome, Singapore 138672, Singapore
FEATURES   Location/Qualifiers
source   1..1873
             /organism="SARS coronavirus Sin0409"
             /mol_type="genomic RNA"
             /db_xref="taxon:266147"
             /country="Singapore"
ORIGIN

Query Match   100.0%; Score 1136; DB 14; Length 1873;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACATTGGCAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   706 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACATTGGCAC 765

```
RESULT 2
AY53476286
LOCUS          AY53476286              1938 bp    RNA     linear   VRL 17-MAR-2004
DEFINITION     SARS coronavirus Sin_WNV, partial sequence.
ACCESSION      AY534767
VERSION        AY534767.1  GI:45384975
KEYWORDS       .
SEGMENT        6 of 6
SOURCE         SARS coronavirus Sin_WNV
  ORGANISM     SARS coronavirus Sin_WNV
               Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
               Coronaviridae; Coronavirus.
REFERENCE      1  (bases 1 to 1938)
  AUTHORS      Lim,P.L., Kurup,A., Gopalakrishna,G., Chan,K.P., Wong,C.W.,
               Ng,L.C., Se-Thoe,S.Y., Oon,L., Bai,X., Stanton,L.W., Ruan,Y.,
               Miller,L.D., Vega,V.B., James,L., Ooi,P.L., Kai,C.S., Olsen,S.J.,
               Ang,B. and Leo,Y.S.
  TITLE        Laboratory-acquired severe acute respiratory syndrome (SARS) -
               Singapore 2003
  JOURNAL      Unpublished
REFERENCE      2  (bases 1 to 1938)
  AUTHORS      Wei,C.L., Lee,C., Lin,S., Thoreau,H., Vega,V.B., Stanton,L.W. and
               Ruan,Y.
  TITLE        Direct Submission
  JOURNAL      Submitted (28-JAN-2004) Genome Institute of Singapore, 60, Biopolis
               Street, 02-01, Genome, Singapore 138672, Singapore
FEATURES             Location/Qualifiers
     source          1..1938
                     /organism="SARS coronavirus Sin_WNV"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:266148"
                     /country="Singapore"
ORIGIN

Query Match              100.0%;  Score 1136;  DB 14;  Length 1938;
Best Local Similarity    100.0%;
Matches 1136;  Conservative     0;  Mismatches     0;  Indels     0;  Gaps     0;
```

```
Qy      1 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACATTGGCAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    732 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACATTGGCAC 791

Qy     61 CCGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGGAACAACATTGCC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    792 CCGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGGAACAACATTGCC 851

Qy    121 AAAAGGCTTCTACGCAGAGGAAGCAGAGGCGCAGTCAAGCCTCTTCTGCTCCTCATC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    852 AAAAGGCTTCTACGCAGAGGAAGCAGAGGCGCAGTCAAGCCTCTTCTGCTCCTCATC 911

Qy    181 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGCAGCAGTAGGGGAAATTCCTGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    912 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGCAGCAGTAGGGGAAATTCCTGC 971

Qy    241 TCGAATGGCTAGCGGAGGTGTGAAACTGCCCTCGCGTATTGCTGTCTAGACAGATTGAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    972 TCGAATGGCTAGCGGAGGTGTGAAACTGCCCTCGCGTATTGCTGTCTAGACAGATTGAA 1031

Qy    301 CCAGCTTGAGAGCATCTAAAACTGCCAAAACAGTACTGCCACAAAACAGTA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1032 CCAGCTTGAGAGCATCTAAAACTGCCAAAACAGTACTGCCACAAAACAGTA 1091

Qy    361 GAAATCTGCTGCTGAGGCATCTAAAACTGCCAAAACAGTACTGCCACAAAACAGTA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1092 GAAATCTGCTGCTGAGGCATCTAAAACTGCCAAAACAGTACTGCCACAAAACAGTA 1151

Qy    421 CAACGTCACTCAAGCATTGGGAGACGTGGTCAGAGAACAAACCCAAGAAATTCGGGGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1152 CAACGTCACTCAAGCATTGGGAGACGTGGTCAGAGAACAAACCCAAGAAATTCGGGGA 1211

Qy    481 CCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGCACAATTTGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1212 CCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGCACAATTTGC 1271

Qy    541 TCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCACACCTTCGGG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1246 TCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCACACCTTCGGG 1305

Qy    601 AACATGGCTGACTTATCATGGAGCCATTAAATTGGATGACAAAGATCCCACAATTCAAAGA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1306 AACATGGCTGACTTATCATGGAGCCATTAAATTGGATGACAAAGATCCCACAATTCAAAGA 1365

Qy    661 CAACGTCATACTGCTGACAAGCACATTCCCACCACAGAGCC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1366 CAACGTCATACTGCTGACAAGCACATTCCCACCACAGAGCC 1425

Qy    721 TAAAAAGGACAAAAGAGAAAAAGAGAAAAAGAAAAAGAGAAAAAGAAAAGAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1426 TAAAAAGGACAAAAGAGAAAAAGAGAAAAAGAAAAAGAA 1485

Qy    781 GCAGCCACTGTGACCTCTTCCTGGCTGACATGGATGATTCTCCAGACAACTTCA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1486 GCAGCCACTGTGACCTCTTCCTGGCTGACATGGATGATTCTCCAGACAACTTCA 1545

Qy    841 AAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAGGCCATAAACACTCATGATGACCAC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1546 AAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAGGCCATAAACACTCATGATGACCAC 1605

Qy    901 ACAAGGCAGATGGCTATGTAAACGTTTTCGTTTACGTAGTCTACTC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1606 ACAAGGCAGATGGCTATGTAAACGTTTTCGTTTACGTAGTCTACTC 1665

Qy    961 TTGTGCAGAATGAATTCTCTGGAGCTTCTGCTAACTAAACAGCACAAGTAGGTTTAATCTC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1666 TTGTGCAGAATGAATTCTCTGGAGCTTCTGCTAACTAAACAGCACAAGTAGGTTTAATCTC 1725

Qy   1021 ACATAGCAATCTTTAATCAATGTGTAACATTAGGGAGGACTTGAAAGAGGCCACCACATTT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1726 ACATAGCAATCTTTAATCAATGTGTAACATTAGGGAGGACTTGAAAGAGGCCACCACATTT 1785

Qy   1081 TCATCGAGGCCACGCGCGGAGTACGATCGAGGGTACAGTGAATAATTGCTAGGGAGAGC 1136
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1786 TCATCGAGGCCACGCGCGGAGTACGATCGAGGGTACAGTGAATAATTGCTAGGGAGAGC 1841
```

```
Qy   541  TCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCACACCTTCGGG  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1272  TCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCACACCTTCGGG 1331

Qy   601  AACATGGCTGACTTATCATGGAGCCATTAAATTGGATGACAAAGATCCACAATTCAAAGA  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1332  AACATGGCTGACTTATCATGGAGCCATTAAATTGGATGACAAAGATCCACAATTCAAAGA 1391

Qy   661  CAACGTCATACTGCTGAACAAGCACATTGACGCACATACAAAACATTCCACCAACAGAGCC  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1392  CAACGTCATACTGCTGAACAAGCACATTGACGCACATACAAAACATTCCACCAACAGAGCC 1451

Qy   721  TAAAAAGGACAAAAGAAAAGACTGATGAGCTCAGCCTTTGCCCGAGACAAAAGAA  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1452  TAAAAAGGACAAAAGAAAAGACTGATGAGCTCAGCCTTTGCCCGAGACAGACAAAAGAA 1511

Qy   781  GCAGCCCACTGTGACTCTTCTTCCTCGCGGCTGACATGGATGATTTCTCCAGACAACTTCA  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1512  GCAGCCCACTGTGACTCTTCTTCCTCGCGGCTGACATGGATGATTTCTCCAGACAACTTCA 1571

Qy   841  AAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAGGCATAAACACTCATGATGACCAC  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1572  AAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAGGCATAAACACTCATGATGACCAC 1631

Qy   901  ACAAGGCAGATGGGCTATGTAAACGTTTTCCCAATTCCGTTTACGGATACAGTCTACTC  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1632  ACAAGGCAGATGGGCTATGTAAACGTTTTCGCAATTCCGTTTACGGATACAGTCTACTC 1691

Qy   961  TTGTGCAGAATGAATTCTCGTAACTAAACAGCACAGTAGTTTAGTTAACTTAATCTC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1692  TTGTGCAGAATGAATTCTCGTAACTAAACAGCACAGTAGTTTAGTTAACTTAATCTC 1751

Qy  1021  ACATAGCAATCTTTAATCAATGTAACATTAGGGAGGACTTGAAAGAGCCACCACCATTT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1752  ACATAGCAATCTTTAATCAATGTAACATTAGGGAGGACTTGAAAGAGCCACCACCATTT 1811

Qy  1081  TCATCGAGGCCACGCGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC 1136
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1812  TCATCGAGGCCACGCGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC 1867
```

RESULT 3
AY290752
LOCUS   AY290752   2810 bp   RNA   linear   VRL 09-FEB-2004
DEFINITION SARS coronavirus ZJ01 isolate ZJ01b uncharacterized protein 6,
            uncharacterized protein 7a, uncharacterized protein 7b,
            uncharacterized protein 8a, uncharacterized protein 8b,
            nucleocapsid protein, uncharacterized protein 9b, and
            uncharacterized protein 9c genes, complete cds.
ACCESSION   AY290752
VERSION   AY290752.2   GI:38735509
KEYWORDS   .
SOURCE   SARS coronavirus ZJ01
  ORGANISM SARS coronavirus ZJ01
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1 (bases 1 to 2810)
  AUTHORS   Li,L., Wang,Z., Lu,Y., Bao,Q., Chen,S., Wu,N., Cheng,S., Weng,J.,
            Zhang,Y., Yan,J., Mei,L., Wang,X., Zhu,H., Yu,Y., Zhang,M., Li,M.,
            Yao,J., Lu,Q., Yao,P., Bo,X., Wo,J., Wang,S. and Hu,S.
  TITLE   Severe acute respiratory syndrome-associated coronavirus genotype
            and its characterization
  JOURNAL   Chin. Med. J. 116 (9), 1288-1292 (2003)
  MEDLINE   22889812
   PUBMED   14527350
REFERENCE   2 (bases 1 to 2810)
  AUTHORS   Wang,Z.G., Li,L.J., Luo,Y., Zhang,J.Y., Wang,M.Y., Cheng,S.Y.,
            Zhang,Y.J., Wang,X.M., Lu,Y.Y., Wu,N.P., Mei,L.L. and Wang,Z.X.
  TITLE   Molecular biological analysis of genotyping and phylogeny of severe
            acute respiratory syndrome associated coronavirus
  JOURNAL   Chin. Med. J. 117 (1), 42-48 (2004)
   PUBMED   14733771

REFERENCE   3 (bases 1 to 2810)
  AUTHORS   Wang,Z., Cheng,S., Zhang,Y., Lu,Y. and Mei,L.
  TITLE   Direct Submission
  JOURNAL   Submitted (03-MAY-2003) Department of Microbiology, Zhejiang Center
            of Disease Prevention and Control, 17 Laozheda Rd., Hangzhou,
            Zhejiang 310009, China
REFERENCE   4 (bases 1 to 2810)
  AUTHORS   Wang,Z., Cheng,S., Zhang,Y., Lu,Y. and Mei,L.
  TITLE   Direct Submission
  JOURNAL   Submitted (05-DEC-2003) Department of Microbiology, Zhejiang Center
            of Disease Prevention and Control, 17 Laozheda Rd., Hangzhou,
            Zhejiang 310009, China
REMARK   Nucleotide and amino acid sequences updated by submitter
COMMENT   On Dec 5, 2003 this sequence version replaced gi:31505969.
FEATURES   Location/Qualifiers
     source   1..2810
                     /organism="SARS coronavirus ZJ01"
                     /mol_type="genomic RNA"
                     /isolate="ZJ01b"
                     /isolation_source="patient with severe acute respiratory
                     syndrome"
                     /db_xref="taxon:230471"
                     /country="China: Hangzhou"
     misc_feature   1..6
                     /note="transcription regulatory core sequence 6"
     CDS   162..353
                     /codon_start=1
                     /product="uncharacterized protein 6"
                     /protein_id="AAR27473.1"
                     /db_xref="GI:38735510"
                     /translation="MFHLVDFQVTIAEILIIIMRTFRIAIWNLDVIISSIVRQLFKPL
                     TKKNYSELDDEEPMELDYP"
     misc_feature   331
                     /note="Xiaohong genotype locus 6"
     misc_feature   355..360
                     /note="transcription regulatory core sequence 7"
     CDS   361..729
                     /codon_start=1
                     /product="uncharacterized protein 7a"
                     /protein_id="AAR27474.1"
                     /db_xref="GI:38735511"
                     /translation="MKIILFLTLIVFTSCELYHYQECVRGTTVLLKEPCPSGTYEGNS
                     PFHPLADNKFALTCSTHFAPACADGTRHTYQLRARSVSPKLFIRQEEVQQELYSPLF
                     LIVAALVFLLCFTIKRKTE"
     CDS   726..860
                     /codon_start=1
                     /product="uncharacterized protein 7b"
                     /protein_id="AAR27475.1"
                     /db_xref="GI:38735512"
                     /translation="MKLLIVLTCISLSCSCICTVVQRCASNKPHVLEDPCKVQH"
     misc_feature   861..866
                     /note="transcription regulatory core sequence 8"
     CDS   867..986
                     /codon_start=1
                     /product="uncharacterized protein 8a"
                     /protein_id="AAR27476.1"
                     /db_xref="GI:38735513"
                     /translation="MKLLIVLTCISLSCSCICTVVQRCASNKPHVLEDPCKVQH"
     misc_feature   915
                     /note="Xiaohong genotype locus 7"
     CDS   952..1206
                     /codon_start=1
                     /product="uncharacterized protein 8b"
                     /protein_id="AAR27477.1"
                     /db_xref="GI:38735514"
                     /translation="MNELTLIDFYLCFLAFLLFLVLIMLIIFWFSLEIQDLEEPCTKV
                     TINCQDPAGGALIARCWYLHEGHQTAAFRDVLVLNKRTN"
     misc_feature   1193..1198
                     /note="transcription regulatory core sequence 9"
     CDS   1208..2476
                     /codon_start=1
                     /product="nucleocapsid protein"

```
Qy    661  CAACGTCATACTGCTGACAAGCACATTGACGACGATACAAAAACATTCCACCAACAGAGCC  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2254  CAACGTCATACTGCTGACAAGCACATTGACGACGATACAAAAACATTCCACCAACAGAGCC  2313

Qy    721  TAAAAAGGACAAAAGAAAAAGACTGATGAAGCTCAGCCTTGCCGCAGAGACAAAAGAA  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2314  TAAAAAGGACAAAAGAAAAAGACTGATGAAGCTCAGCCTTGCCGCAGAGACAAAAGAA  2373

Qy    781  GCAGCCCACTGTGACTCTTCTTCCGCGGCTGACATGGATGATTTCTCCGACAACTTCA  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2374  GCAGCCCACTGTGACTCTTCTTCCGCGGCTGACATGGATGATTTCTCCGACAACTTCA  2433

Qy    841  AAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAGGCATAAACACTCATGATGACCAC  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2434  AAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAGGCATAAACACTCATGATGACCAC  2493

Qy    901  ACAAGGCAGATGGGCTATGTAAACGTTTCGCAATTCCGTTTACGATACATAGTCTACTC  960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2494  ACAAGGCAGATGGGCTATGTAAACGTTTCGCAATTCCGTTTACGATACATAGTCTACTC  2553

Qy    961  TTGTGTCAGAATGAATTCTCGTAACTAAACAGCACAAGTAGGTTTAGTTAACTTTAATCTC 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2554  TTGTGTCAGAATGAATTCTCGTAACTAAACAGCACAAGTAGGTTTAGTTAACTTTAATCTC 2613

Qy   1021  ACATTAGCAATCTTTAATCAATGTGTAACATTAGGGGACACTTGAAAGAGCCACCACATT 1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2614  ACATTAGCAATCTTTAATCAATGTGTAACATTAGGGGACACTTGAAAGAGCCACCACATT 2673

Qy   1081  TCATCGAGCCACCGGAGTACGACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC 1136
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2674  TCATCGAGCCACCGGAGTACGACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC 2729
```

```
RESULT 4
AY304487
LOCUS        AY304487                8581 bp     RNA     linear      VRL  05-NOV-2003
DEFINITION   SARS coronavirus SZ13, partial genome.
ACCESSION    AY304487
VERSION      AY304487.1  GI:34482138
KEYWORDS     .
SOURCE       SARS coronavirus SZ13
  ORGANISM   SARS coronavirus SZ13
             Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
             Coronaviridae; Coronavirus.
REFERENCE    1  (bases 1 to 8581)
  AUTHORS    Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
             Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,
             Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and
             Poon,L.L.M.
  TITLE      Isolation and characterization of viruses related to the SARS
             coronavirus from animals in southern China
  JOURNAL    Science 302 (5643), 276-278 (2003)
  MEDLINE    22913660
   PUBMED    12958366
REFERENCE    2  (bases 1 to 8581)
  AUTHORS    Guan,Y. and Zheng,B.J.
  TITLE      Direct Submission
  JOURNAL    Submitted (26-MAY-2003) Microbiology, The University of Hong Kong,
             University Pathology Building, Queen Mary Hospital, Pokfulam Road,
             Hong Kong, China
FEATURES             Location/Qualifiers
     source          1..8581
                     /organism="SARS coronavirus SZ13"
                     /mol_type="genomic RNA"
                     /isolate="SZ13"
                     /db_xref="taxon:231514"
                     /country="Hong Kong"
ORIGIN
Query Match             100.0%;  Score 1136;  DB 14;  Length 8581;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches  1136;          Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
                     /protein_id="AAP44772.1"
                     /db_xref="GI:31505970"
                     /translation="MSDNGPQSNQRSAPRITFGGPTDSTDNNQNGGRNGARPKQRRPQ
                     GLPNNTASWFTALTQHGKEELFPPRQGVPINTNSGPDDQIGYYRRATRRVRGGDGKM
                     KELSPRWYFYYLGTGPEASLPYGANKEGIVWVATEGALNTPKDHIGTRNPNNNAATVL
                     QLPQGTTLPKGFYAEGSRGGSQASSRSSSRGNSRNSTPGGSRGNSPARMASGGGET
                     ALALLLDRLNQLESKVSGKGQQQQGQTVTKKSAAEASKKPRQKRTATKQYNVTQAFG
                     RRGPEQTQGNFGDQDLIRQGTDYKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTY
                     HGAIKLDDKDQFKDNVILLNKHIDAYKTFPPTEPKKDKKKKTDEAQPLPQRQKKQPT
                     VTLLPADMDDFSRQLQNSMSGASADSTQA"
CDS              1218..1514
                     /codon_start=1
                     /product="uncharacterized protein 9b"
                     /protein_id="AAP69659.1"
                     /db_xref="GI:31747858"
                     /translation="MDPNQTNVVPPALHLVDPQIQLTITRMEDAMGQGQNSADPKVYP
                     IILRLGSQLSLSMARRNLDSLEARAFQSTPIVQMTKLATTEELPDFEVVTAK"
CDS              1671..1883
                     /codon_start=1
                     /product="uncharacterized protein 9c"
                     /protein_id="AAP69660.1"
                     /db_xref="GI:31747859"
                     /translation="MLPPCYNFLKEQHCQKASTQREAEAAVKPLLAPHHVVAVIQEIQ
                     LLAAVGEILLLEWLAEVVKLPSRYCC"
ORIGIN

Query Match             100.0%;  Score 1136;  DB 14;  Length 2810;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches  1136;          Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1  AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACATTGGCAC    60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1594  AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACATTGGCAC  1653

Qy     61  CCGGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGGAACAACATTGCC   120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1654  CCGGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGGAACAACATTGCC  1713

Qy    121  AAAAGGCTTCTACGCCAGAGGAGGCAGAGGCGGAGAGCAGTGTCAAGCCTTCTTCGCTCCTCATC   180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1714  AAAAGGCTTCTACGCCAGAGGAGGCAGAGGCGGAGAGCAGTGTCAAGCCTTCTTCGCTCCTCATC  1773

Qy    181  ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGGCAGCAGTAGGGGAAATTCTCCTGC   240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1774  ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGGCAGCAGTAGGGGAAATTCTCCTGC  1833

Qy    241  TCGAATGGCTAGCCGGAGTGGTGAAACTGCCCTTCGCCCTATTGCTGCTAGACAGATTGAA   300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1834  TCGAATGGCTAGCCGGAGTGGTGAAACTGCCCTTCGCCCTATTGCTGCTAGACAGATTGAA  1893

Qy    301  CCAGCTTGAGAGCAAAGTTTCTGGTAAAGGCCAACAACAACAGGCCAAACTGTCACTAA   360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1894  CCAGCTTGAGAGCAAAGTTTCTGGTAAAGGCCAACAACAACAGGCCAAACTGTCACTAA  1953

Qy    361  GAAATCTGCTGCTGAGGCATCTAAAAAGCCTCGCCAAAAACGTACTGCCACAAAACAGTA   420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1954  GAAATCTGCTGCTGAGGCATCTAAAAAGCCTCGCCAAAAACGTACTGCCACAAAACAGTA  2013

Qy    421  CAACGTCACTCAAGCATTGGGAGACGTGGTCAGAGAACAAAGCCCAGGAAATTTCGGGGA   480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2014  CAACGTCACTCAAGCATTGGGAGACGTGGTCAGAGAACAAAGCCCAGGAAATTTCGGGGA  2073

Qy    481  CCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCCCAAATTGCACAAATTGC   540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2074  CCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCCCAAATTGCACAAATTGC  2133

Qy    541  TCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCACACCTTCGGG   600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2134  TCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCACACCTTCGGG  2193

Qy    601  AACATGGCTGACTTATCATGAGCCATTAAATTGGATGACAAAGATCCACAATTCAAAGA   660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2194  AACATGGCTGACTTATCATGAGCCATTAAATTGGATGACAAAGATCCACAATTCAAAGA  2253
```

us-10-764-075-1.rge

Tue May 17 15:13:33 2005

Right column:

```
Qy   1081 TCATCGAGGCCACGCGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC 1136
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8450 TCATCGAGGCCACGCGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC 8505

RESULT 5
AY304493
LOCUS       AY304493    11010 bp    RNA    linear    VRL 05-NOV-2003
DEFINITION  SARS coronavirus HKU-65806, partial genome.
ACCESSION   AY304493
VERSION     AY304493.1   GI:34482144
KEYWORDS    .
SOURCE      SARS coronavirus HKU-65806
  ORGANISM  SARS coronavirus HKU-65806
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 11010)
  AUTHORS   Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
            Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,
            Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and
            Poon,L.L.M.
  TITLE     Isolation and characterization of viruses related to the SARS
            coronavirus from animals in southern China
  JOURNAL   Science 302 (5643), 276-278 (2003)
  MEDLINE   22913660
   PUBMED   12958366
REFERENCE   2  (bases 1 to 11010)
  AUTHORS   Guan,Y. and Zheng,B.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAY-2003) Microbiology, The University of Hong Kong,
            University Pathology Building, Queen Mary Hospital, Pokfulam Road,
            Hong Kong, China
FEATURES            Location/Qualifiers
     source          1..11010
                     /organism="SARS coronavirus HKU-65806"
                     /mol_type="genomic RNA"
                     /isolate="HKU-65806"
                     /db_xref="taxon:231520"
                     /country="Hong Kong"
ORIGIN

Query Match             100.0%;  Score 1136;  DB 14;  Length 11010;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1136;    Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 AGGCATCGTATGGTTGCAACTTGAGGAGCCTTGAATACACCCAAAGACCACATTGGCAC 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9781 AGGCATCGTATGGTTGCAACTTGAGGAGCCTTGAATACACCCAAAGACCACATTGGCAC 9840

Qy     61 CCGCAATCCTAATAACAATGCTGCTACAACTTCCTCAAGGAACAACATTGCC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9841 CCGCAATCCTAATAACAATGCTGCTACAACTTCCTCAAGGAACAACATTGCC 9900

Qy    121 AAAAGGCTTCTACGCAGAGGAAGCAGAGGGCGGCAGAGGGCGG 180
          |||||||||||||||||||||||||||||||||||||||||||
Db   9901 AAAAGGCTTCTACGCAGAGGAAGCAGAGGGCGGCAGAGGGCGG 9960

Qy    181 ACGTAGTCGCGGTAATTCAACTCCTGCAGTAGGGGAAATTCTCCTGC 240
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   9961 ACGTAGTCGCGGTAATTCAACTCCTGCAGTAGGGGAAATTCTCCTGC 10020

Qy    241 TCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTCGCGGCTATTGCTGTCTAGACAGATTGAA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10021 TCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTCGCGGCTATTGCTGTCTAGACAGATTGAA 10080

Qy    301 CCAGCTTGAGAGCAAAGTTTCGGTAAAGCCAACAACAAGGCCAAACTGTCACTAA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10081 CCAGCTTGAGAGCAAAGTTTCGGTAAAGCCAACAACAAGGCCAAACTGTCACTAA 10140

Qy    361 GAAATCTGCTGCTGAGGCATCTAAAAAGCCTCGCCAAAAAACAGTA 420
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  10141 GAAATCTGCTGCTGAGGCATCTAAAAAGCCTCGCCAAAAAACAGTA 10200
```

Left column:

```
Qy      1 AGGCATCGTATGGTTGCAACTTGAGGAGCCTTGAATACACCCAAAGACCACATTGGCAC 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7370 AGGCATCGTATGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACATTGGCAC 7429

Qy     61 CCGCAATCCTAATAACAATGCTGCTACAACTTCCTCAAGGAACAACATTGCC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7430 CCGCAATCCTAATAACAATGCTGCTACAACTTCCTCAAGGAACAACATTGCC 7489

Qy    121 AAAAGGCTTCTACGCAGAGGAAGCAGAGGGCGGCAGAGGGCGG 180
          |||||||||||||||||||||||||||||||||||||||||||
Db   7490 AAAAGGCTTCTACGCAGAGGAAGCAGAGGGCGGCAGAGGGCGG 7549

Qy    181 ACGTAGTCGCGGTAATTCAACTCCTGCAGTAGGGGAAATTCTCCTGC 240
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   7550 ACGTAGTCGCGGTAATTCAACTCCTGCAGTAGGGGAAATTCTCCTGC 7609

Qy    241 TCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTCGCGGCTATTGCTGTCTAGACAGATTGAA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7610 TCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTCGCGGCTATTGCTGTCTAGACAGATTGAA 7669

Qy    301 CCAGCTTGAGAGCAAAGTTTCGGTAAAGGCCAACAACAAGGCCAAACTGTCACTAA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7670 CCAGCTTGAGAGCAAAGTTTCGGTAAAGGCCAACAACAAGGCCAAACTGTCACTAA 7729

Qy    361 GAAATCTGCTGCTGAGGCATCTAAAAAGCCTCGCCAAAAAACAGTA 420
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   7730 GAAATCTGCTGCTGAGGCATCTAAAAAGCCTCGCCAAAAAACAGTA 7789

Qy    421 CAACGTCACTCAAGCATTTGGGAGACGTGGTCCAGAACAAACCCAAGGAAATTTCGGGGA 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7790 CAACGTCACTCAAGCATTTGGGAGACGTGGTCCAGAACAAACCCAAGGAAATTTCGGGGA 7849

Qy    481 CCAAGACCTAATCAGACAAGGAACTGATTACAACATTGGCCGCAAATTGCACAAATTGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7850 CCAAGACCTAATCAGACAAGGAACTGATTACAACATTGGCCGCAAATTGCACAAATTGC 7909

Qy    541 TCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCACACCTTCGGG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7910 TCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCACACCTTCGGG 7969

Qy    601 AACATGGCTGACTTATCATGAGCCCATTAAATTGATGACAAAGATCCACAATTCAAAGA 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7970 AACATGGCTGACTTATCATGAGCCCATTAAATTGGATGACAAAGATCCACAATTCAAAGA 8029

Qy    661 CAACGTTCATACTGCTGAACAAGCACATTGACGCTTCGAGCTTCCTGTCTACAACTCCACCGAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8030 CAACGTTCATACTGCTGAACAAGCACATTGACGCTTCGAGCTTCCTGTCTACAACTCCACCGAGC 8089

Qy    721 TAAAAGGACAAAAGAAAGAAAAAAGACTGATGAGCTCAGCCTTTGCCGCAGAGACAAAGAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8090 TAAAAGGACAAAAGAAAGAAAAAAGACTGATGAGCTCAGCCTTTGCCGCAGAGACAAAGAA 8149

Qy    781 GCAGCCCACTGTGACTCTTCTTCCTGCGGCTGACATGGATGATTGATTCTCCAGACAACTTCA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8150 GCAGCCCACTGTGACTCTTCTTCCTGCGGCTGACATGGATGATTGATTCTCCAGACAACTTCA 8209

Qy    841 AAATTCCATGAGTGGAGCTTCTCGTGATTCAACTCATAAAACACTCATGATGACCAC 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8210 AAATTCCATGAGTGGAGCTTCTCGTGATTCAACTCATAAAACACTCATGATGACCAC 8269

Qy    901 ACAAGGCAGAGATGGGCTATGTAAACGTTTTCCGTTTACGATACATAGTCTACTC 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8270 ACAAGGCAGAGATGGGCTATGTAAACGTTTTCCGTTTACGATACATAGTCTACTC 8329

Qy    961 TTGTGTCAGAATGAATTCTCCTGAACTAAACAGCACAAGTAGGTTTAGTTAACTTTAATCTC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8330 TTGTGTCAGAATGAATTCTCCTGAACTAAACAGCACAAGTAGGTTTAGTTAACTTTAATCTC 8389

Qy   1021 ACATAGCAATCTTTAATCAATGTAACATTAGGGAGGACTTGAAAGAGCCACCATTT 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8390 ACATAGCAATCTTTAATCAATGTAACATTAGGGAGGACTTGAAAGAGCCACCATTT 8449
```

```
JOURNAL      Submitted (27-MAY-2003) Microbiology, The University of Hong Kong,
             University Pathology Building, Queen Mary Hospital, Pokfulam Road,
             Hong Kong, China
FEATURES             Location/Qualifiers
     source          1..11010
                     /organism="SARS coronavirus HKU-66078"
                     /mol_type="genomic RNA"
                     /isolate="HKU-66078"
                     /db_xref="taxon:231521"
                     /country="Hong Kong"
ORIGIN

Query Match         100.0%;  Score 1136;  DB 14;  Length 11010;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1136;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 AGGCATCGTATGGGTTGAACTGAGGGAGCCTTGAATACACCCAAAGACCACATTGGCAC   60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     9781 AGGCATCGTATGGGTTGAACTGAGGGAGCCTTGAATACACCCAAAGACCACATTGGCAC 9840

Qy       61 CCGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGGAACAACATTGCC  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     9841 CCGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGGAACAACATTGCC 9900

Qy      121 AAAAGGCTTCACCGCAGAGGAAGCAGAGCGGCAGTCAAGCCTTCTTCGCGTCCTCATC  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     9901 AAAAGGCTTCACCGCAGAGGAAGCAGAGCGGCAGTCAAGCCTTCTTCGCGTCCTCATC 9960

Qy      181 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGGCAGCAGTAGGGGAAATTCTCCTGC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     9961 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGGCAGCAGTAGGGGAAATTCTCCTGC 10020

Qy      241 TCGAATGGCTAGCCGGAGGTGGTGAAACTGCCCCCGCGCCATTGCTGCTAGCAGATTGAA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    10021 TCGAATGGCTAGCCGGAGGTGGTGAAACTGCCCCCGCGCCATTGCTGCTAGCAGATTGAA 10080

Qy      301 CCAGCTTGAGACAAAGTTTCTGGTAAAGCCAACAACAACAAGGCCAAACTGTCACTAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    10081 CCAGCTTGAGACAAAGTTTCTGGTAAAGCCAACAACAACAAGGCCAAACTGTCACTAA 10140

Qy      361 GAAATCTGCTGCGAGCATCTAAAAAGCCTCGCCACAAAACGTACTGCCACAAAACAGTA  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    10141 GAAATCTGCTGCGAGCATCTAAAAAGCCTCGCCACAAAACGTACTGCCACAAAACAGTA 10200

Qy      421 CAACGTCACTCAAGCATTTGGGACGACTGTGTCCAGAACAAACCAAGGAAATTTCGGGGA  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    10201 CAACGTCACTCAAGCATTTGGGACGACTGTGTCCAGAACAAACCAAGGAAATTTCGGGGA 10260

Qy      481 CCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCCAAATTGCACAATTGC  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    10261 CCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCCAAATTGCACAATTGC 10320

Qy      541 TCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCACACCTTCGG  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    10321 TCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCACACCTTCGG 10380

Qy      601 AACATGGCTGACTTATCATGAGCCATTAAATTGGATGACAAAGATCCCAACAGAGCC  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    10381 AACATGGCTGACTTATCATGAGCCATTAAATTGGATGACAAAGATCCCAACAGAGCC 10440

Qy      661 CAACGTCATCACTGCTGAACAAGCACATTGACGCATACAAAAACATTCCCACAGAGC  720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    10441 CAACGTCATCACTGCTGAACAAGCACATTGACGCATACAAAAACATTCCCACAGAGC 10500

Qy      721 TAAAAAGGACAAAAAGAAAAAGAAAAGACTGATGAAGCTCAGCCTTTGCCGCAGAGACAAAAGAA  780
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    10501 TAAAAAGGACAAAAAGAAAAAGAAAAGACTGATGAAGCTCAGCCTTTGCCGCAGAGACAAAAGAA 10560

Qy      781 GCAGCCCACTGTGACTCTTCTTCTTCGCGGCTGACATGGATGATTTCTCCAGACAACTTCA  840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    10561 GCAGCCCACTGTGACTCTTCTTCTTCGCGGCTGACATGGATGATTTCTCCAGACAACTTCA 10620

Qy      841 AAATTCCATGAGTGCTGAGTTCAACTCAGGCATAAACACTCATGATGACCAC  900
```

```
Qy      421 CAACGTCACTCAAGCATTTGGGAGACGTGGTCCAGAACAAACCAAGGAAATTTCGGGGA  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    10201 CAACGTCACTCAAGCATTTGGGAGACGTGGTCCAGAACAAACCAAGGAAATTTCGGGGA 10260

Qy      481 CCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCCAAATTGCACAATTGC  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    10261 CCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCCAAATTGCACAATTGC 10320

Qy      541 TCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCACACCTTCGG  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    10321 TCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCACACCTTCGG 10380

Qy      601 AACATGGCTGACTTATCATGAGCCATTAAATTGGATGACAAAGATCCACCAATTCAAAGA  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    10381 AACATGGCTGACTTATCATGAGCCATTAAATTGGATGACAAAGATCCACCAATTCAAAGA 10440

Qy      661 CAACGTCATACTGCTGAACAAGCACATTGACGCATACAAAAACATTCCCACCAACAGAGCC  720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    10441 CAACGTCATACTGCTGAACAAGCACATTGACGCATACAAAAACATTCCCACCAACAGAGCC 10500

Qy      721 TAAAAAGGACAAAAAGAAAAAGAAAGACTCAGCCTTTGCGCAGAGACAAAAGAA  780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    10501 TAAAAAGGACAAAAAGAAAAAGAAAGACTCAGCCTTTGCGCAGAGACAAAAGAA 10560

Qy      781 GCAGCCCACTGTGACTCTTCTTCTTCGCGGCTGACATGGATGATTTCTCCAGACAACTTCA  840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    10561 GCAGCCCACTGTGACTCTTCTTCTTCGCGGCTGACATGGATGATTTCTCCAGACAACTTCA 10620

Qy      841 AAATTCCATGAGTGCTGAGTTCAACTCAGGCATAAACACTCATGATGACCAC  900
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    10621 AAATTCCATGAGTGCTGAGTTCAACTCAGGCATAAACACTCATGATGACCAC 10680

Qy      901 ACAAGGCAGAGATGGCTATGTGTAAACGTTTTCGCAATTCCGTTTACGATACATAGTCTACTC  960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    10681 ACAAGGCAGAGATGGCTATGTGTAAACGTTTTCGCAATTCCGTTTACGATACATAGTCTACTC 10740

Qy      961 TTGTGCAGAATGAATTCTCGTACTAAACAGCACAGTAGGTTTAGTTAACTTAATCTC 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    10741 TTGTGCAGAATGAATTCTCGTACTAAACAGCACAGTAGGTTTAGTTAACTTAATCTC 10800

Qy     1021 ACATAGCAATCTTTAATCAATGTGTAACATTAGGGAGGACTGAAAGAGCCACCACATTT 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    10801 ACATAGCAATCTTTAATCAATGTGTAACATTAGGGAGGACTGAAAGAGCCACCACATTT 10860

Qy     1081 TCATCGAGGCCACCGGGAGTACGACGCGGAGGGTACAGTCAGTGAATAATGCTAGGGAGAGC 1136
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    10861 TCATCGAGGCCACCGGGAGTACGACGCGGAGGGTACAGTCAGTGAATAATGCTAGGGAGAGC 10916
```

```
RESULT 6
AY304494
LOCUS       AY304494      11010 bp    RNA      linear      VRL 05-NOV-2003
DEFINITION  SARS coronavirus HKU-66078, partial genome.
ACCESSION   AY304494
VERSION     AY304494.1  GI:34482145
KEYWORDS    .
SOURCE      SARS coronavirus HKU-66078
  ORGANISM  SARS coronavirus HKU-66078
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 11010)
  AUTHORS   Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
            Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,
            Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and
            Poon,L.L.M.
  TITLE     Isolation and characterization of viruses related to the SARS
            coronavirus from animals in southern China
  JOURNAL   Science 302 (5643), 276-278 (2003)
  MEDLINE   22913660
   PUBMED   12958366
REFERENCE   2  (bases 1 to 11010)
  AUTHORS   Guan,Y. and Zheng,B.J.
  TITLE     Direct Submission
```

Left column — sequence alignment (continued)

```
Db  10621 AAATTCCATGAGTGGAGCTTCGTGCTGATTCAACTCAGGCATAAAACACTCATGATGACCAC 10680
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    901 ACAAGGCAGATGGGCTATGTGAAACGTTTCGCAATTCCGTTTACGTAACATAGTCTACTC 960
Db  10681 ACAAGGCAGATGGGCTATGTGAAACGTTTCGCAATTCCGTTTACGTAACATAGTCTACTC 10740
Qy    961 TTGTGCAGAATGAATTCTCGTAACTAAACAGCACAAGTAGGTTTAGTTAACTTTAATCTC 1020
Db  10741 TTGTGCAGAATGAATTCTCGTAACTAAACAGCACAAGTAGGTTTAGTTAACTTTAATCTC 10800
Qy   1021 ACATAGCAATCTTTAATCAATGTGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 1080
Db  10801 ACATAGCAATCTTTAATCAATGTGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 10860
Qy   1081 TCATCGAGGCCACCGCCGGATACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC 1136
Db  10861 TCATCGAGGCCACCGCCGGATACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC 10916
```

RESULT 7
AY304492

LOCUS   AY304492   13471 bp   RNA   linear   VRL 05-NOV-2003
DEFINITION SARS coronavirus HKU-36871, partial genome.
ACCESSION   AY304492
VERSION   AY304492.1 GI:33482143
KEYWORDS   .
SOURCE   SARS coronavirus HKU-36871
  ORGANISM SARS coronavirus HKU-36871
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1 (bases 1 to 13471)
  AUTHORS   Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
            Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,
            Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and
            Poon,L.L.M.
  TITLE   Isolation and characterization of viruses related to the SARS
            coronavirus from animals in southern China
  JOURNAL   Science 302 (5643), 276-278 (2003)
  MEDLINE   22913660
   PUBMED   12958366
REFERENCE   2 (bases 1 to 13471)
  AUTHORS   Guan,Y. and Zheng,B.J.
  TITLE   Direct Submission
  JOURNAL   Submitted (27-MAY-2003) Microbiology, The University of Hong Kong,
            University Pathology Building, Queen Mary Hospital, Pokfulam Road,
            Hong Kong, China
FEATURES   Location/Qualifiers
     source   1..13471
                     /organism="SARS coronavirus HKU-36871"
                     /mol_type="genomic RNA"
                     /isolate="HKU-36871"
                     /db_xref="taxon:231519"
                     /country="Hong Kong"
ORIGIN

Query Match   100.0%;   Score 1136;   DB 14;   Length 13471;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1136;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy      1 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACATTGGCAC 60
Db  12242 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACATTGGCAC 12301
Qy     61 CCGCAATCCTAATAACAATCTGCCACCGTGCTACACTTCCTCAAGGACGACATTGCC 120
Db  12302 CCGCAATCCTAATAACAATCTGCCACCGTGCTACACTTCCTCAAGGACGACATTGCC 12361
Qy    121 AAAAGGCTTCTACCGCAGAGGGAGCAGAGAGCGCGAGTCCTCTTCGCTTCCTCATC 180
Db  12362 AAAAGGCTTCTACCGCAGAGGGAGCAGAGAGCGCGAGTCCTCTTCGCTTCCTCATC 12421
Qy    181 ACGTAGTCGCGGTAATTCAACTCCTCGGCAGTAGGGGAAATTCTCCTGC 240
```

Right column — sequence alignment (continued)

```
Db  12422 AGTAGTCGCGGTAATTCAACTCCTCGGCAGTAGGGGAAATTCTCCTGC 12481
Qy    241 TCGAATGGCTAGCGGCGAGTGGGTGAAACTGCCCTCGCCTATTGCTAGCAGACGATTGAA 300
Db  12482 TCGAATGGCTAGCGGCGAGTGGGTGAAACTGCCCTCGCCTATTGCTAGCAGACGATTGAA 12541
Qy    301 CCAGCTTGAGAGCAAAGTTTCTGGTAAAGGCCAACAACAAGCAGCCAAACTGTCACTAA 360
Db  12542 CCAGCTTGAGAGCAAAGTTTCTGGTAAAGGCCAACAACAAGCAGCCAAACTGTCACTAA 12601
Qy    361 GAAATCTGCTGTGAGGCATCTAAAAGCCTCGCCACAAAACTACTGCCACAAAACAGTA 420
Db  12602 GAAATCTGCTGTGAGGCATCTAAAAGCCTCGCCACAAAACTACTGCCACAAAACAGTA 12661
Qy    421 CAACGTCACTCAAGCATTGGAGACGTGGTCCAGAAGACAAACCCAGGAAATTGCACAA 480
Db  12662 CAACGTCACTCAAGCATTGGAGACGTGGTCCAGAAGACAAACCCAGGAAATTGCACAA 12721
Qy    481 CCAAGACCTAATCAGACAAGGAACTAGATTACAAAACATTGGCCGCAAATTGCACAATTTGC 540
Db  12722 CCAAGACCTAATCAGACAAGGAACTAGATTACAAAACATTGGCCCAAATTGCACAATTTGC 12781
Qy    541 TCCAAGTGCCTCTGCCATTCTTTGGAATGTCACGCATGGAAGTCACACCTTCGGG 600
Db  12782 TCCAAGTGCCTCTGCCATTCTTTGGAATGTCACGCATGGAAGTCACACCTTCGGG 12841
Qy    601 AACATGGCTGACTTATCATGGAGCCATTAAATTGATGACAAAGATCCACAATTCAAAGA 660
Db  12842 AACATGGCTGACTTATCATGGAGCCATTAAATTGATGACAAAGATCCACAATTCAAAGA 12901
Qy    661 CAACGTCATACTGCTGACAAGCAGCATCACATTGAGCGATCTGACATGAGAGCC 720
Db  12902 CAACGTCATACTGCTGACAAGCAGCATCACATTGAGCGATCTGACATGAGAGCC 12961
Qy    721 TAAAAAGGACAAAAAGAAAAGAAAAAGAAAAAGACTGAGCTCAGCCTTTGCCGAGAGACAAAAGAA 780
Db  12962 TAAAAAGGACAAAAAGAAAAGAAAAAGAAAAAGACTGAGCTCAGCCTTTGCCGAGAGACAAAAGAA 13021
Qy    781 GCAGCCCACTGACTCTTCTTCCTGCGCGTGACATGGATGATTTCTCCGAGACAACTTCA 840
Db  13022 GCAGCCCACTGACTCTTCTTCCTGCGCGTGACATGGATGATTTCTCCGAGACAACTTCA 13081
Qy    841 AAATTCCATGAGTGGAGCTTCGTGCTGATTCAACTCAGGCATAAACACTCATGATGACCAC 900
Db  13082 AAATTCCATGAGTGGAGCTTCGTGCTGATTCAACTCAGGCATAAACACTCATGATGACCAC 13141
Qy    901 ACAAGGCAGATGGGCTATGTGAAACGTTTCGCAATTCCGTTTACGTAACATAGTCTACTC 960
Db  13142 ACAAGGCAGATGGGCTATGTGAAACGTTTCGCAATTCCGTTTACGTAACATAGTCTACTC 13201
Qy    961 TTGTGCAGAATGAATTCTCGTAACTAAACAGCACAAGTAGGTTTAGTTAACTTTAATCTC 1020
Db  13202 TTGTGCAGAATGAATTCTCGTAACTAAACAGCACAAGTAGGTTTAGTTAACTTTAATCTC 13261
Qy   1021 ACATAGCAATCTTTAATCAATGTGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 1080
Db  13262 ACATAGCAATCTTTAATCAATGTGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 13321
Qy   1081 TCATCGAGGCCACCGCCGGATACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC 1136
Db  13322 TCATCGAGGCCACCGCCGGATACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC 13377
```

RESULT 8
AY463060

LOCUS   AY463060   29013 bp   RNA   linear   VRL 05-JAN-2004
DEFINITION SARS coronavirus ShanghaiQXC2, complete genome.
ACCESSION   AY463060
VERSION   AY463060.1 GI:40457448
KEYWORDS   .
SOURCE   SARS coronavirus ShanghaiQXC2
  ORGANISM SARS coronavirus ShanghaiQXC2

viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE   1 (bases 1 to 29013)
AUTHORS   Yuan,Z., Zhang,X., Hu,Y., Lan,S., Zhou,Z., Wang,H. and Wen,Y.
TITLE   Analysis of SARS coronavirus genome in Shanghai isolates
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 29013)
AUTHORS   Yuan,Z., Zhang,X., Hu,Y., Lan,S., Zhou,Z., Wang,H. and Wen,Y.
TITLE   Direct Submission
JOURNAL   Submitted (11-NOV-2003) Key Lab of Medical Molecular Virology,
            Shanghai Medical College, Fudan University, 138 Yi Xue Yuan Road,
            Shanghai 200032, P.R.China
FEATURES   Location/Qualifiers
     source   1..29013
                     /organism="SARS coronavirus ShanghaiQXC2"
                     /mol_type="genomic RNA"
                     /strain="ShanghaiQXC2"
                     /db_xref="taxon:258508"
     CDS   join(186..12734,12734..20827)
                     /note="ribosomal slippage"
                     /codon_start=1
                     /product="orf1ab"
                     /protein_id="AAR86774.1"
                     /db_xref="GI:40457449"
                     /translation="MESLVLGVNEKTHVQLSLPVLQVRDVLVRGFGDSVEEALSEARE
                     HLKNGTCGLVELEKGVLPQLEQPYVFIKRSDALSTNHGHKVVELVAEMDGIQYGRSGI
                     TLGVLVPHVAGETPIAYRNVLLRKNGNKGAGGHSYGIDLKSYDLGDELGTDPIEYGVG
                     WNTKHGSGALRELTRELNGGAVTRYDNNFCGPDGVPLDCIKDFLARAGKSMCTLSEQ
                     LDYIESKRGVYCCRDHEHEIAWFTERSDKSYEHQTPFEIKSAKKFDTFKGECPKFVFP
                     LNSKVKVIQPRVEKKKTEGFMGRIRSVPVASPQECNNMHLSTLMKCNHCDEVSWQTC
                     DFLKATCEHCGTENLVIEGPTTCGYLPTNAVVKMPCPACQDPEIGPEHSVADYHNHSN
                     IETRLRKGGRTRCFGGCVFAYVGCYNKRAYWVPRASADIGSGHTGITGDNVETLNEDL
                     LEILSRERVNINIVGDFHLNEVAIILASFSASTSAFIDTIKSLDYKSFKTIVESGCHN
                     YKVTKGKPVKGAWNICQQRSVLTPLCGFPSQAAGVIRSIFARTLDAANHSIPDLQRAA
                     VTILDGISEQSLRLVDAMVYTSDLLTNSVIIMAYVTGGLVQQTSQWLSNLLGTTVEKL
                     RPIFEWIEAKLSAGVEFLKDAWEILKFLITCGEDTVWEVGDVKVNRITFELDERVDKVLNE
                     KCSVYTVESGTEVTEFACVVAEAVVKTLQPVSDLLTNMGIDLDEWSVATFYLFDDAGE
                     ENFSSRMYCSFYPPDEEEEDDAECEEEEIDETCEHEYGTEDDYQGLPLEFGASAETVR
                     VEEEEEDWLDDTTEQSEIEPEPEPTPEEPVNQFTGYLKLTNNGAVAIKDVFVLPSPAV
                     NPMVIVNAANIHLKHGGGVAGALNKATNGAMQKESDDYIKLNGPLTVGGSCLLSGHNL
                     AKKCLHVVGPNLNAGEDIQLLKAAYENFNSQDILLAPLLSAEIGDLTDHPLLSQVCVQT
                     VRTQVYIAVNDKALYEQVVMDYLDNLKPRVEAPKQEEPPNTEDSKTEEKSVVQKAYQKV
                     KPKIKACIDEVTTLEETKFLTNKLLLFADINGKLYHDSQNMLRGEDMSFLEEDAPYM
                     VGDVITSGDITCVVIPSKKAGGTTEMLSRALLKKVPVDEYITTYPGQCAGYTLEEAKTI
                     ALKKCKSAFYVLPSEAPNAKEEILGTVSWNLAEHLAHEAEARKLMPICMDVRAIMATI
                     QRKYKGVKIQEGIVDYGVRFFFYTSKEPVASIITKLNSLNEPLVTMPIGYVTHGFNLE
                     EAARCMRSLKAPAVVSVSSPDAVTTNGYLTRSSSKKTSEEHFVETVSLAGSYRDWSYSG
                     QRTELGVEFLKRGDKIYHTLESPVEFHLDGEVLSLDKLKSLLSLREVKTIKVFTTVD
                     NTNLHTQLVDMSMTYGQQFGPTYLDGADVTKIKPHVNHEGKTFFVLPSDDTLRSEAFE
                     YYHTLDESFLGRYMSALNHTKKWKFPQVGGLTSIKWADNNCYLSSVLLALQQLEVKFN
                     APALQEAYYRARAGDAANFCALILAYSNKTVKLHKDTDVRETMTHLQHANLESAKRVLN
                     VCVCKHCGQKTTLTGVEAVMYMGTLSYDNLKTGVSIPCVCGRDATQYLVQESSFVMM
                     SAPPAGYKLQQGTFLCANEYTGNYQCGHYTHITAKETLYRIDGAHLTKMSEYKGPVTD
                     VFYKETSYTTTIKPVSYKLDGVTYTEIPKLGVYKKDNAYVENTSYITIKPNELSLA
                     LGLKTIATHGIAAINSVPWSKILAYVKPFLGQAATTSNCAKRLAQRVFNNYMPVVFT
                     LLLSICLGSLICVTAAFGYFISNLMWFIISIVQMAPVSAMVRMIYFASFYIVWKSYVHIMDGCTS
                     STCMCYKCNHATRVECTTIVNGMKRSFYVYANGGRGFCKTHNWNCLNCDTFCTGSTF
                     ISDEVARDLSLQFKRPINPTDQSSYIVDSVAVKNGALPHYPDKAGQKTYERHPLSHFV
                     NLDNLRANNTKGSLPINVIVFDGKSKCDESASKSASVYYSQLMCQPILLLDQALVSDV
                     GDSTEVSVKMFDAYVDTFSATFSVPMEKLKALVATAHSELAKGVALDGVLSTFVSAAR
                     QGVVDTVDTKDVIECLKLSHHSDLEVTGDSCNFMLTYNKVENMTPRDLGACIDCNA
                     RHINAQVAKSHNVSLIWNVKDYMSLSEQLRKQIRSAAKKNIPFRLTCATTRQVVNVI
                     TTKISLKGGKIVSTCFKLMLKATLLCVLAALVCYIVMPVHTLSIHDGYTNEIIGYKAI
                     QDGVTRDIISTDDCFANKHAGFDAWFSQRGGSYKNDKSCPVVAAIITREIGFIVPGLP
                     GTVLRAINGDFLHFLPRVFSAVGNICYTPSKLIEYSDFATSACVLAAECTIFKDAMGK
                     PVPYCYDTNLLEGSISYSELRPDTRYVLMDGSIIQFNTYLEGSVRVVTTFDAEYCRH"

GTCERSEVGICLSTSGRWVLNNEHYRALSGVFCGVDAMNLIANIFTPLVQPVGALDAS
ASVVAGGIIALLVTCAAYYFMKFRRAFGEYNHVVAANALLFLMSFTILCLVPAYSFLP
GVYSVFYLYLTFYFTNDVSFLAHLQWFAMFSPIVPFWITAIYVFCISLKHCHWFFNNY
LRKRVMFNGVTFSTFEEAALCTFLLNKEMYLKLRSETLLPTQNYRYLALYNRYKYFS
GVLDTTSYREAACCHLAKALNDFSNSGADVLYQPPQTSITSAVLQSGFRKMAPPSGKV
EGCMVQVTCGTTTLNGLWLDDTVYCPRHVICTAEDMLNPNYEDLLIRKSNHSFLVQAG
NVQLRVIGHSMQNCLLRLKVDTSNPKTPKYKFVRIQPGQTFSVLACYNGSPSGVYQCA
MRPNHTIKGSFLNGSCGSVGFNIDYDCVSFCYMHHMELPTGVHAGTDLEGKFYGPFVD
RQTAQAAGTDTTITLNVLAWLYAAVINGDRWFLNRFTTLDNFNLVAMKYNYEPLTQD
HVDILGPLSAQTGIAVLDMCAALKELLQNGMNGRTILGSTILEDEFTPFDVVRQCSGV
TFQGKFKKIVKGTHHWMLLTFLTSLLVLVQSTQWSLFFFVYENAFLPFTLGIMAIAAC
AMLLVKHKHAFLCLFLLPSLATVAYFNMVYMPASWVMRIMTWLELADTSLSGYRLKDC
VMYASALVLLIMTARTVYDDGAARRVWTLMNVTLVYKYVYGNALDQAISMWALIVSA
TSNYSGVVTTIMFLARAIVFVCVEYYPLLFITGNTLQCIMLVYCFLGYCCCYFGLFC
LLNRYFRLTLGVYDYLVSTQERFRYMNSQGLLPPKSSIDAFKLNIKLLGIGGKPCIKVA
TVQSKMSDVKCTSVVLLSVLQQLRVESSSKLWAQCVQLHNDILLAKDTTEAFEKMVSL
LSVLLSMQGAVDINRLCEEVLDNRATLQAIASEFSSLPSYAAYATAQEAYEQAVANGD
SEVVLKKLKKSLNVAKSEFDRDAAMQRKLEKMADQAMTQMYKQARSEDKRAKVTSGMT
TMLFTMRKLDNDALNNIINNARDGCVPLNIIPLTTAAKLMVVVPDYGTYKNTCDGNT
FTYASALWEIQQVDADSKIVQLSEINDNSPNLAWPLIVTALRANSAAKLQNNELSP
VALRQMSCAAGTTQTACTDDNALAYYNNSKGGRFVLALLSDHQDLKWARFPKSDGTGT
IYTELEPPCRFVTDTPKGPKMKYLYFIKGLNNLNRGMVGLSLAATVRLQAGNATEVPA
NSTVLSPCAFAVDPAKAYKDYLASGGQPITNCVKMLCTHTGTGQAITVTPEANMDQGS
FGGASCCLYCRCHIDHPNPKGFCDLKGKYVQIPTTCANDPVGFTLRNTVCTVCGMWKG
YGCSCDQLREPLMQSADASTFFKRVCGVSAARLTPCGTGTSTDVVYRAPDIYNEKVAG
FAKFLKTNCCRFQEKDEEGNLLDSYFVVKRHTMSNYQHEETIYNLIKDCPAVAVHDFF
KFRVDGDMVPHISRQRLTKYTMADLVALRHFPEGNCDTLKEILVTYNCCDDDYFNKK
DWDFVENPDILRVYANLGERVRQSLLKTVQFCDAMRDAGIVGVLTLDNQDLNGNWYD
FGDFVQVAPGCGVPIVDSYYSLLMPILTLTRALAAESHMDADLAKPLIKWDLLKYDFT
EERLCLFDRYFKYWDQTYHPNCINCLDDRCILRACNFLFSTVFPPTSFGPLVRKIF
VDGVPFVVSTGYHFRELGVVHNQTGELAIANGNALSLSVAADPAMHAASGNLLLDKR
TTCFSVAALTNNVAFQTVKPGNFNKDFYDFAVSKGFFKEGSSVELKHFFFAQDGNAAI
SDYDYYRNLPTMCDIRQLLFVEVVDKYFDCYDGGCINANQVIVNNLDKSAGFPFNK
WGKARLYYDSMSYEDQDALFAYTKRNVIPTITQMNLKYAISAKNRARTVAGVSICSTM
TNRQFHQKLLKSIAATRGATVVIGTSKFYGGWHNMLKTVYSDVETPHLMGWDYPKCDR
AMPNMLRIMASLVLARKHNTCCNLSHRFYRLANECAQVLSEMVMCGGSLYVKPGGTSS
GDATTAYANSVFNICQAVTANVNALLSTDGNKIADKYVRNLQHRLYECLYRNRDVDHE
FVDEFYAYLRKHFSMMLSDDAVVCYNSNYAAQGLVASIKNFKAVLYYQNNVFMSEAK
CWTETDLTKGPHEFCSQHTMLVKQGDDYVYLPYPDPSRILGAGCFVDDIVKTDGTLMI
ERFVSLAIDAYPLTQHYDQEYADHFVDLYQYIRKLHDELTGHMLDMYSVMLTNDNTSR
YWEPEFYEAMYTPHTVLQAVGACVLCNSQTSLRCGACIRRPFLCCKCCYDHVISTSHK
LVLSVNPYVCNAPGCDVTDVTNGDYIANTCTERLKLEHKILSYCFKLSYGIATVRE
GSDNVTDFNAIATCDWTNAGDYILANTCTERLKLEHKIISQDLSVISKVVKVTGLH
VLSDRELHLSWEVGKPRPPLNRNYVFTGYRVTKNSKVQIGEYTFEKGDYGDAVVYRGT
TTYKLNVGDYFVLTSHTVMPLSAPTLVPQEHYVRITGLYPTLNISDEFSSNVANYQKV
GMQKYSTLQGPPGTGKSHFAIGLALYYPSARIVYTACSHAAVDALCEKALKYLPIDKC
SRIIPARARVECFDKFKVNSTLEQYVFCTVNALPETTADIVVFDEISMATNYDLSVVN
ARLRAKHVYIGDPAQLPAPRTLLTKGTLEPEYFNSVCRLMKTIGPDMFLGTCRRCPA
EIVDTVSALVYDNKLKAHKEKSAQCFKMFYKGVITHDVSSAINRPQIGVVREFLTRNP
AWRKAVFISPYNSQNAVASKILGLPTQVTMDIAKKPTESACSSLTVLFDGRVEGQ
VAITRAKIGILCIMSDRDLYDKLQFTSLEIPRRNVATLQAENVTGLFKDCSKIITGLH
PTQAPTHLSVDIKFKTEGLCVDIPGIPKDMTYRRLISMMGFKMNYQVNGYPNMFITRE
EAIRHVRAWIGFDVEGCHATRDAVGTNLPLQLGFSTGVNLVAVPTGYVDTENNTEFTR
VNAKPPPGDQFKHLIPLMYKGLPWNVVRIVQMLSDTLKGLSDRVVFVLWAHGFELT
SMKYFVKIGPERTCCLCDKRATCFSTSSDTYACWNHSVGFDYVYNPFMIDVQQWGFTG
NLQSNHDQHCQVGNAHVASCDAIMTRCLAVHECFVKRVDWSVEYPIIGDELRVNSAC
RKVQHMVVKSALLADKFPVLHDIGNPKAIKCVPQAEVEWKFYDAQPCSDKAYKIEELF
YSYATHHDKFTDGVCLFWNCNVDRYPANAIVCRPFDTRVLSNLNLPGCDGGSLYVNKHA
FHTPAFDKSAFTNLKQLPFFYSDSPCESHGKQVVSDIDYVPLKSATCITRCNLGGAV
CRHHANEYRQYLDAYNMISAGFSLWIYKQFDTYNLWNTFTRLQSLENVAYNVVNKGH
FDGHAGEAPVSIINAVTIWDYKREAPAHVSTIGVCTMTDIAKKPTESACSSLTVLFDGRVEGQ
NNLGVDIAANTVIWDYKREAPAHVSTIGVCTMTDIAKKPTESACSSLTVLFDGRVEGQ
VDLFRNARNGVLITEGSVKGLTIGESKGPAQASANVGLTIGESKGPAQASANVGLIIQQ
LPETYFTQSRDLEDFRPRSQMETDFLELAMDEFIQRYKLEGYAFEHIVYGDFSHQLG
GLHLMIGLAKRSQDSPLKLEDFIPMDSTVKNYFITDAQTGSSKCVRSVIDLLLDDFVE
IIKSQDLSVISKVVKVTGLHAEISFMLWCKDGHVETFYPKLQASQAWQPGVAMPNLYK
MQRMLLEKCDLQNYGENAVIPKGIMMNVAKYTQLCQYLNTLTLAVPYNMRVIHFGAGS
DKGVAPGTAVLRQWLPTGTLLVDSDLNDFVSDADSTLIGDCATVHTANKWDLIISDMY
DPRTKHVIKENDSKEGFFTYLCGFIKQKLALGGSIAVKITEHSWNAQLYKLMGHFSWW
TAFVTNVNASSEAFLIGANYLGKPKEQIDGYTMHANYIFWRNTNPIQLSSYSLFDMS
KPPLKLRGTAVMSLKENQINDMIYSLLEKGRLIIRENNRVVVSSDILVNN"

                     20834..24601
                     /codon_start=1
CDS

```
/product="spike"
/protein_id="AAR86775.1"
/db_xref="GI:40457450"
/translation="MFIFLLFLTLTSGSDLDRCTTFDDVQAPNYTQHTSSMRGVYYPD
EIFRSDTLYLTQDLFLPFYSNVTGFHTINHTFGNPVIPFKDGIYFAATEKSNVVRGWV
FGSTMNNKSQSVIIINNSTNVVIRACNFELCDNPFFAVSKPMGTQTHTMIFDNAFNCT
FEYISDAFSLDVSEKSGNFKHLREFVFKNKDGFLYVYKGYQPIDVVRDLPSGFNTLKP
IFKLPLGINITNFRAILTAFSPAQDIWGTSAAAYFVGYLIKPTIFMLKYDENGTITDAV
DCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSGDVVRFPNITNLCPFGEVFNATKF
PSVYAWERKISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCFSNVYADSFVVKGD
DVRQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRP
FERDISNVPFSPDGKPCTPPALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPA
TVCGPKLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPK
TSEILDISPCAFGGVSVITPGTNASSEVAVLYQDVNCTNVSAAIHADQLTPAWRIYST
GNNVFQTQAGCLIGAEHVDTSYECDIPIGAGICASYHTVSLLRSTSQKSIVAYTMSLG
ADSSIAYSNNTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLQYGS
FCTQLNRALSGIAAEQDRNTREVFAQVKQMYKTPTLKYFGGFNFSQILPDPLKPTKRS
FIEDLLFNKVTLADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYT
AALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAIS
QIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAE
VQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVFDCGKGYH
LMSFPQAAPHGVVFLHVTYVPSQERNFTTAPAICHEGKAYFPREGVFVNGTSWFITQ
RNFFSPQITTDNTFVSGNCDVVIGIINNTVYDPLQPELDSFKGELDKYFKNHTSPDV
DLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIA
GLIAIVMVTILLCCMTSCCSCLKGACSCGSCCKFDEDDSEPVLKGVKLHYT"
CDS   24610. .25434
      /codon_start=1
      /product="unknown"
      /protein_id="AAR86776.1"
      /db_xref="GI:40457451"
      /translation="MDLFMRFFLLGSITAQPVKIDNASPASTAHATATIPLQASLPFG
WLVIGVAFLAVFQSATKIIALNKRWQLALYKGFQPICNLLLLFVTIYSHLLLVAAGME
AQFLYLYALIYFLQCINACRIIMRCWLCWKCKSKNPLLYDANYFVCWHTHNYDYCIPY
NSVTDTVVTEGDGISTPKLKEDYIQGYSEDRHSGVKDYVVVHGYFTEVYYQLESTQ
ITTNTGIENATFFIFNKLVKDPPNVQIHTIDGSSGVANPAMDPIYDEPTTTTSVPL"
CDS   25031. .25495
      /codon_start=1
      /product="unknown"
      /protein_id="AAR86777.1"
      /db_xref="GI:40457452"
      /translation="MMPPTLFAGTHIMTTVYHITVSQIQLSLLKVTAFQHQNSKKTT
KLVVILRIGTQVLKTMSLYMAISPKFTSLSHKLLQTIVLKMLHSSSLTSLLKTHRM
CKYTQSTALQELLIQQWIQFMMSRRRLLACLCKHKKVSTNLCTHSFRKKQVR"
CDS   25459. .25689
      /codon_start=1
      /product="E"
      /protein_id="AAR86778.1"
```

```
Query Match      100.0%;  Score 1136;  DB 14;  Length 29013;
Best Local Similarity 100.0%;  Pred. No. 0;
Matches 1136;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy        1 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCAAAGACCACCATTGGCAC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    27848 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCAAAGACCACCATTGGCAC 27907

Qy       61 CCGCAATCCTAATAACAATGCTGCCACCGTGCCACCGTGCTTCCTCAAGGAGGACAACATTGGCC 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    27908 CCGCAATCCTAATAACAATGCTGCCACCGTGCTTCCTCAAGGAGGACAACATTGGCC 27967

Qy      121 AAAAGGCTTCTACCGCAGAGGGAAGCAGAGCGGCAGTCAAGCCTTCTTCTCGCAGCCTCATC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    27968 AAAAGGCTTCTACCGCAGAGGGAAGCAGAGCGGCAGTCAAGCCTTCTTCTCGCAGCCTCATC 28027

Qy      181 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGCAGCAGTAGGGGAAATTCCTGC 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    28028 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGCAGCAGTAGGGGAAATTCCTGC 28087

Qy      241 TCGAATGGCTAGCCGGAGGTGTGAAACTGCCCTGCCCGTCGTATTGCTGCTAGCACAGATTGAA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    28088 TCGAATGGCTAGCCGGAGGTGTGAAACTGCCCTGCCCGTCGTATTGCTGCTAGCACAGATTGAA 28147

Qy      301 CCAGCTTGAGAGCAAAGTTTCTGGTAAAGGCCAAACTGTCACTAA 360
```

```
Db    28148 CCAGCTTGAGAGCAAAGTTTCTGGTAAAGGCCAAACTGTCACTAA 28207

Qy      361 GAAATCTGCTGCTGAGGCATCTGAGGGAGACGTGGTCACAGAGACTACTGCCAAAACAGTA 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    28208 GAAATCTGCTGCTGAGGCATCTGAGGGAGACGTGGTCACAGAGACTACTGCCAAAACAGTA 28267

Qy      421 CAACGTCACTCAAGCATTGGGAGACGTGGTCACAGAGACAACCCAAGGAAATTTCGGGGA 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    28268 CAACGTCACTCAAGCATTGGGAGACGTGGTCACAGAGACAACCCAAGGAAATTTCGGGGA 28327

Qy      481 CCAAGACCTAATCAGACAAGGAACTGATTACAAAACATTGGCCGCCAAATTGCACAATTGC 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    28328 CCAAGACCTAATCAGACAAGGAACTGATTACAAAACATTGGCCGCCAAATTGCACAATTGC 28387

Qy      541 TCCAAGTGCCTCGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCACACCTTCGGG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    28388 TCCAAGTGCCTCGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCACACCTTCGGG 28447

Qy      601 AACATGGCTACTATCGAGCAATTGGATGACAAAGATCCACAATTCAAAGA 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    28448 AACATGGCTACTATCGAGCAATTGGATGACAAAGATCCACAATTCAAAGA 28507

Qy      661 CAACGTCATACTGCTGACAAGCACATTCCCACACAGAGCC 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    28508 CAACGTCATACTGCTGACAAGCACATTCCCACACAGAGCC 28567

Qy      721 TAAAAGGACAAAAGAAAAGACTGATGAGCTCAGCCTTTGCCGCAGAGACAAAAGAA 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    28568 TAAAAGGACAAAAGAAAAGACTGATGAGCTCAGCCTTTGCCGCAGAGACAAAAGAA 28627

Qy      781 GCAGCCCACTGACTCTTCCTCGCGGCTGACATGATGATTTCCTCGACAACTTCA 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    28628 GCAGCCCACTGACTCTTCCTCGCGGCTGACATGATGATTTCCTCGACAACTTCA 28687

Qy      841 AAATTCCATGAGTGGAGCTTCGTCGTGATTCAACTCAGCATAAACACTCATGATGACCAC 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    28688 AAATTCCATGAGTGGAGCTTCGTCGTGATTCAACTCAGCATAAACACTCATGATGACCAC 28747

Qy      901 ACAAGGCAGATGGGCTATGTGAAACGTTTCGCAATTCCGTTTACGATACATAGTCTACTC 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    28748 ACAAGGCAGATGGGCTATGTGAAACGTTTCGCAATTCCGTTTACGATACATAGTCTACTC 28807

Qy      961 TTGTGCAGAATGAATTCCGTAACTAAACAGCACAAGTAGGTTTAGTTAACTTTAATCTC 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    28808 TTGTGCAGAATGAATTCCGTAACTAAACAGCACAAGTAGGTTTAGTTAACTTTAATCTC 28867

Qy     1021 ACATAGCAATCTTTAATCAATGTGTAACAATGTGTAACATTAGGGAGGACTTGAAAGAGCCACACATTT 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    28868 ACATAGCAATCTTTAATCAATGTGTAACAATGTGTAACATTAGGGAGGACTTGAAAGAGCCACACATTT 28927

Qy     1081 TCATCGAGGCCACGCCGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGGAGC 1136
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    28928 TCATCGAGGCCACGCCGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGGAGC 28983
```

```
RESULT 9
AY394999
LOCUS       AY394999      29350 bp    RNA    linear   VRL 29-JAN-2004
DEFINITION  SARS coronavirus LC2, complete genome.
ACCESSION   AY394999
VERSION     AY394999.1  GI:37624342
KEYWORDS    .
SOURCE      SARS coronavirus LC2
  ORGANISM  SARS coronavirus LC2
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 29350)
  AUTHORS   The SARS epidemiology consortium of Guangdong
  CONSRTM   From independent foci of epidemic outbreak to large genomic
  TITLE     alteration in late phase viruses: evolution of the SARS-coronavirus
  JOURNAL   Unpublished
```

```
Qy   841  AAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAGCATAAACACTCATGATGACCAC  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28945  AAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAGCATAAACACTCATGATGACCAC  29004

Qy   901  ACAAGGCAGATGGGCTATGTAAACGTTTTCGCAATTCCGTTACGATACATAGTCTACTC  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29005  ACAAGGCAGATGGGCTATGTAAACGTTTTCGCAATTCCGTTACGATACATAGTCTACTC  29064

Qy   961  TTGTGCAGAATGAATTCTGTAACTAAACAGCACACAAGTAGTTTAGTTAGTTAATCTC  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29065  TTGTGCAGAATGAATTCTGTAACTAAACAGCACACAAGTAGTTTAGTTAGTTAATCTC  29124

Qy  1021  ACATAGCAATCTTAATCAATGTGAACATTAGGGAGGACTTGAAAGAGCCACCACATT  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29125  ACATAGCAATCTTAATCAATGTGAACATTAGGGAGGACTTGAAAGAGCCACCACATT  29184

Qy  1081  TCATCGAGGCCACGCGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC  1136
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29185  TCATCGAGGCCACGCGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC  29240
```

RESULT 10
AY395000
LOCUS   AY395000   29350 bp   RNA   linear   VRL 29-JAN-2004
DEFINITION   SARS coronavirus LC3, complete genome.
ACCESSION   AY395000
VERSION   AY395000.1 GI:37624343
KEYWORDS   .
SOURCE   SARS coronavirus LC3
  ORGANISM   SARS coronavirus LC3
             Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
             Coronaviridae; Coronavirus.
REFERENCE   1 (bases 1 to 29350)
  AUTHORS
  CONSRTM   The SARS epidemiology consortium of Guangdong
  TITLE   From independent foci of epidemic outbreak to large genomic
             alteration in late phase viruses: evolution of the SARS-coronavirus
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 29350)
  AUTHORS   .
  CONSRTM   The SARS epidemiology consortium of Guangdong
  TITLE   Direct Submission
  JOURNAL   Submitted (19-SEP-2003) Guangdong, China
FEATURES   Location/Qualifiers
     source   1..29350
                     /organism="SARS coronavirus LC3"
                     /mol_type="genomic RNA"
                     /isolate="LC3"
                     /db_xref="taxon:249085"
ORIGIN

Query Match   100.0%;   Score 1136;   DB 14;   Length 29350;
Best Local Similarity 100.0%;   Pred. No. 0;
Matches 1136;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

```
Qy     1  AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACATTGGCAC   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28105  AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACATTGGCAC  28164

Qy    61  CCGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGGAGACAACATTGCC  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28165  CCGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGGAGACAACATTGCC  28224

Qy   121  AAAAGGCTTCTACGCAGAGGGAAGCAGAGGCGGCAGTCAAGCCTCTTCTCCGCCTCCTCATC  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28225  AAAAGGCTTCTACGCAGAGGGAAGCAGAGGCGGCAGTCAAGCCTCTTCTCCGCCTCCTCATC  28284

Qy   181  ACGTAGTCGCGGTAATTCAAGAAATTCAAGAAATTGCTGCTAGAGTAGGGGAAATTCTCCTGC  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28285  ACGTAGTCGCGGTAATTCAAGAAATTCAAGAAATTGCTGCTAGAGTAGGGGAAATTCTCCTGC  28344

Qy   241  TCGAATGGCTAGCCGAGGTGGTGAAACTGCCCTCGCCGCTATTGCTGCTAGACAGATTGAA  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

TITLE   From independent foci of epidemic outbreak to large genomic
               alteration in late phase viruses: evolution of the SARS-coronavirus
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 29350)
AUTHORS
CONSRTM   The SARS epidemiology consortium of Guangdong
TITLE   Direct Submission
JOURNAL   Submitted (19-SEP-2003) Guangdong, China
FEATURES   Location/Qualifiers
   source   1..29350
               /organism="SARS coronavirus LC4"
               /mol_type="genomic RNA"
               /isolate="LC4"
               /db_xref="taxon:249086"

ORIGIN
Query Match   100.0%;   Score 1136;   DB 14;   Length 29350;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1136;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTGAATACACCCAAAGACCACATTGGCAC   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28105 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTGAATACACCCAAAGACCACATTGGCAC 28164

Qy   61 CCGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGGAACAACATTGCC   120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28165 CCGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGGAACAACATTGCC 28224

Qy   121 AAAAGGCTTCTACGCAGAGGAAGCAGAGGGCGGCAGTCAAGCTCTTCGCTTCCTCATC   180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28225 AAAAGGCTTCTACGCAGAGGAAGCAGAGGGCGGCAGTCAAGCTCTTCGCTTCCTCATC 28284

Qy   181 ACGTAGTCGCGTAATTCAAGAAATTCAACTCCTGGCAGCAGTAGGGGAAATTCTCCTGC   240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28285 ACGTAGTCGCGTAATTCAAGAAATTCAACTCCTGGCAGCAGTAGGGGAAATTCTCCTGC 28344

Qy   241 TCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTGCGCGGGCTATTGCTGTCGTAGACAGATTGAA   300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28345 TCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTGCGCGGGCTATTGCTGTCGTAGACAGATTGAA 28404

Qy   301 CCAGCTTGAGAGCAAAGTTTCGTGGTAAAGGCCAACAAGGCCAAACTGTCACTAA   360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28405 CCAGCTTGAGAGCAAAGTTTCGTGGTAAAGGCCAACAAGGCCAAACTGTCACTAA 28464

Qy   361 GAAATCTGCTGCTGAGGCATCTAAAAAGCCTGCCAAAAAACGTACTGCCACAAAAACAGTA   420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28465 GAAATCTGCTGCTGAGGCATCTAAAAAGCCTGCCAAAAAACGTACTGCCACAAAAACAGTA 28524

Qy   421 CAACGTCACTCAAGCATTTGGGAGACGTGGTCCAGAACAACCCAAGGAAATTTCGGGGA   480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28525 CAACGTCACTCAAGCATTTGGGAGACGTGGTCCAGAACAACCCAAGGAAATTTCGGGGA 28584

Qy   481 CCAAGACCTAATCAGACAAGAACTGATTACAAAACATTGGCCGCAAATTGCACAATTTGC   540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28585 CCAAGACCTAATCAGACAAGAACTGATTACAAAACATTGGCCGCAAATTGCACAATTTGC 28644

Qy   541 TCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAGTCACACCTTCGGG   600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28645 TCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAGTCACACCTTCGGG 28704

Qy   601 AACATGGCTGACTTATCATCGAGCCATTAAATTGGATGACAAAGATCCAAGATCCAAGA   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28705 AACATGGCTGACTTATCATCGAGCCATTAAATTGGATGACAAAGATCCAAGATCCAAGA 28764

Qy   661 CAACGTCATACTGCTGAACAAGCAGCATACAAAAACATTCCACCAACAGAGCC   720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28765 CAACGTCATACTGCTGAACAAGCAGCATACAAAAACATTCCACCAACAGAGCC 28824

Qy   721 TAAAAAGGACAAAAGAAAAAGAAAAGGACTGATGAAGCTCAGCCTTTTGCCGCCGAGAGACAAAAGAA   780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28825 TAAAAAGGACAAAAGAAAAAGAAAAGGACTGATGAAGCTCAGCCTTTTGCCGCCGAGAGACAAAAGAA 28884

Qy   781 GCAGCCCACTGTGACTTCTTCCTCGCGGCTGACATGGATGATTTCTTCCAGACAACTTCA   840

Db 28345 TCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTGCGCGGGCTATTGCTGTCGTAGACAGATTGAA 28404
Qy   301 CCAGCTTGAGAGCAAAGTTTCGTGGTAAAGGCCAACAAGGCCAAACTGTCACTAA   360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28405 CCAGCTTGAGAGCAAAGTTTCGTGGTAAAGGCCAACAAGGCCAAACTGTCACTAA 28464
Qy   361 GAAATCTGCTGCTGAGGCATCTAAAAAGCCTGCCAAAAAACGTACTGCCACAAAAACAGTA   420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28465 GAAATCTGCTGCTGAGGCATCTAAAAAGCCTGCCAAAAAACGTACTGCCACAAAAACAGTA 28524
Qy   421 CAACGTCACTCAAGCATTTGGGAGACGTGGTCCAGAACAACCCAAGGAAATTTCGGGGA   480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28525 CAACGTCACTCAAGCATTTGGGAGACGTGGTCCAGAACAACCCAAGGAAATTTCGGGGA 28584
Qy   481 CCAAGACCTAATCAGACAAGAACTGATTACAAAACATTGGCCGCAAATTGCACAATTTGC   540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28585 CCAAGACCTAATCAGACAAGAACTGATTACAAAACATTGGCCGCAAATTGCACAATTTGC 28644
Qy   541 TCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAGTCACACCTTCGGG   600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28645 TCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAGTCACACCTTCGGG 28704
Qy   601 AACATGGCTGACTTATCATCGAGCCATTAAATTGGATGACAAAGATCCAAGATCCAAGA   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28705 AACATGGCTGACTTATCATCGAGCCATTAAATTGGATGACAAAGATCCAAGATCCAAGA 28764
Qy   661 CAACGTCATACTGCTGAACAAGCAGCATACAAAAACATTCCACCAACAGAGCC   720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28765 CAACGTCATACTGCTGAACAAGCAGCATACAAAAACATTCCACCAACAGAGCC 28824
Qy   721 TAAAAAGGACAAAAGAAAAAGAAAAGGACTGATGAAGCTCAGCCTTTTGCCGCCGAGAGACAAAAGAA   780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28825 TAAAAAGGACAAAAGAAAAAGAAAAGGACTGATGAAGCTCAGCCTTTTGCCGCCGAGAGACAAAAGAA 28884
Qy   781 GCAGCCCACTGTGACTTCTTCCTCGCGGCTGACATGGATGATTTCTTCCAGACAACTTCA   840

RESULT 11
AY395001
LOCUS   AY395001   29350 bp   RNA   linear   VRL 29-JAN-2004
DEFINITION SARS coronavirus LC4, complete genome.
ACCESSION   AY395001
VERSION   AY395001.1 GI:37624344
KEYWORDS   .
SOURCE   SARS coronavirus LC4
  ORGANISM SARS coronavirus LC4
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1 (bases 1 to 29350)
  AUTHORS
  CONSRTM   The SARS epidemiology consortium of Guangdong

Qy 241 TCGAATGGCTAGCGGCGGAGGTGGTGAAACTGCCCTCGCGCTCATTGTGCTAGACAGATTGAA 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28345 TCGAATGGCTAGCGGCGGAGGTGGTGAAACTGCCCTCGCGCTCATTGTGCTAGACAGATTGAA 28404

Qy 301 CCAGCTTGAGAGCAAAGTTTCTGGTAAAGGCCAACAACAACAAGGCCAAACTGTCACTAA 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28405 CCAGCTTGAGAGCAAAGTTTCTGGTAAAGGCCAACAACAACAAGGCCAAACTGTCACTAA 28464

Qy 361 GAAATCTGCTGCTGAGGCATCTAAAAAGCCTCGCCACAAAACGTACTGCCACAAAACAGTA 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28465 GAAATCTGCTGCTGAGGCATCTAAAAAGCCTCGCCACAAAACGTACTGCCACAAAACAGTA 28524

Qy 421 CAACGTCACTCAAGCATTGGGACGACGTGGTCCAGAACAACCCAAGGAAATTCGGGGA 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28525 CAACGTCACTCAAGCATTGGGACGACGTGGTCCAGAACAACCCAAGGAAATTCGGGGA 28584

Qy 481 CCAAGACCTAATCAGACAAGAAGAAGACGAACTGATTACAACATTGCCGCCGCAAATTGCACAATTTGC 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28585 CCAAGACCTAATCAGACAAGAAGAAGACGAACTGATTACAACATTGCCGCCGCAAATTGCACAATTTGC 28644

Qy 541 TCCAAGTGCCTCTGCATTCTTGGAATGTCACGCGCATTGGGAATGTCACGCGCATTGGGAGAGCGC 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28645 TCCAAGTGCCTCTGCATTCTTGGAATGTCACGCGCATTGGGAATGTCACGCGCATTGGGAGAGCGC 28704

Qy 601 AACATGGCTGACTTATCATGGAGCCATTAAAATTGGATGACAAAGATCCACAAATTCAAAGA 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28705 AACATGGCTGACTTATCATGGAGCCATTAAAATTGGATGACAAAGATCCACAAATTCAAAGA 28764

Qy 661 CAACGTCATACTGCTGAACAAGCACATTGACGATACAAACAAACATTCCCACCAACAGAGCC 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28765 CAACGTCATACTGCTGAACAAGCACATTGACGATACAAACAAACATTCCCACCAACAGAGCC 28824

Qy 721 TAAAAGGACAAAAAGAAAAGACTGATGAAGCTCAGCCTTTGCCGCAGAGACAAAAGAA 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28825 TAAAAGGACAAAAAGAAAAGACTGATGAAGCTCAGCCTTTGCCGCAGAGACAAAAGAA 28884

Qy 781 GCAGCCCACTGTGACTCTTCTTCCTCGCGGCTGACATGGATGATTTCTCCAGACAACTTCA 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28885 GCAGCCCACTGTGACTCTTCTTCCTCGCGGCTGACATGGATGATTTCTCCAGACAACTTCA 28944

Qy 841 AAATTCCATGCAGATGGCTATGTGGAGCTTCGTGATTCAACTCAACTCATGATGACCAC 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28945 AAATTCCATGCAGATGGCTATGTGGAGCTTCGTGATTCAACTCAACTCATGATGACCAC 29004

Qy 901 ACAAGGCAGATGGGCTATGTGAAACGTTTTCGCAATTCCGTTTACGATACATAGTCTACTC 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29005 ACAAGGCAGATGGGCTATGTGAAACGTTTTCGCAATTCCGTTTACGATACATAGTCTACTC 29064

Qy 961 TTGTGCAGAATGAATTCTGTAACATAGTGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29065 TTGTGCAGAATGAATTCTGTAACATAGTGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 29124

Qy 1021 ACATAGCAATCTTAATCAATGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29125 ACATAGCAATCTTAATCAATGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 29184

Qy 1081 TCATCGAGGCCACGCGCCACGCGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC 1136
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29185 TCATCGAGGCCACGCGCCACGCGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC 29240

RESULT 13
AY394985
LOCUS   AY394985   29530 bp   RNA   linear   VRL 29-JAN-2004
DEFINITION   SARS coronavirus HSZ-Bb, complete genome.
ACCESSION   AY394985
VERSION   AY394985.1   GI:37624328
KEYWORDS   .
SOURCE   SARS coronavirus HSZ-Bb
  ORGANISM   SARS coronavirus HSZ-Bb
             Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
             Coronaviridae; Coronavirus.

Db 28885 GCAGCCCACTGTGACTCTTCTTCCTCGCGGCTGACATGGATGATTTCTCCAGACAACTTCA 28944
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 841 AAATTCCATGCAGATGGCTATGTGGAGCTTCGTGATTCAACTCAACTCATGATGACCAC 900

Db 28945 AAATTCCATGCAGATGGCTATGTGGAGCTTCGTGATTCAACTCAACTCATGATGACCAC 29004
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 901 ACAAGGCAGATGGGCTATGTGAAACGTTTTCGCAATTCCGTTTACGATACATAGTCTACTC 960

Db 29005 ACAAGGCAGATGGGCTATGTGAAACGTTTTCGCAATTCCGTTTACGATACATAGTCTACTC 29064
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 961 TTGTGCAGAATGAATTCTGTAACATAGTGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 1020

Db 29065 TTGTGCAGAATGAATTCTGTAACATAGTGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 29124
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1021 ACATAGCAATCTTAATCAATGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 1080

Db 29125 ACATAGCAATCTTAATCAATGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 29184
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1081 TCATCGAGGCCACGCGCCACGCGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC 1136

Db 29185 TCATCGAGGCCACGCGCCACGCGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC 29240

RESULT 12
AY395002
LOCUS   AY395002   29350 bp   RNA   linear   VRL 29-JAN-2004
DEFINITION   SARS coronavirus LC5, complete genome.
ACCESSION   AY395002
VERSION   AY395002.1   GI:37624345
KEYWORDS   .
SOURCE   SARS coronavirus LC5
  ORGANISM   SARS coronavirus LC5
             Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
             Coronaviridae; Coronavirus.
REFERENCE   1 (bases 1 to 29350)
  AUTHORS   The SARS epidemiology consortium of Guangdong
  CONSRTM   From independent foci of epidemic outbreak to large genomic
  TITLE   alteration in late phase viruses: evolution of the SARS-coronavirus
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 29350)
  AUTHORS   The SARS epidemiology consortium of Guangdong
  CONSRTM   Direct Submission
  TITLE   Submitted (19-SEP-2003) Guangdong, China
  JOURNAL   Guangdong, China
FEATURES   Location/Qualifiers
     source   1..29350
                     /organism="SARS coronavirus LC5"
                     /mol_type="genomic RNA"
                     /isolate="LC5"
                     /db_xref="taxon:249087"
ORIGIN
Query Match   100.0%;   Score 1136;   DB 14;   Length 29350;
Best Local Similarity 100.0%;   Pred. No. 0;
Matches 1136;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy   1 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACCATTGGCAC 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28105 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACCATTGGCAC 28164

Qy   61 CCGCAATCCTAATAACAATGCTGCCACCGTACCACGTGCTACTACTTCCTCCTGCTACAACAACATTGCC 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28165 CCGCAATCCTAATAACAATGCTGCCACCGTACCACGTGCTACTACTTCCTCCAAGAACAACATTGCC 28224

Qy 121 AAAAGGCTTCTACCGCAGAGGGAAGCAGAGGCGGCAGTCGAGTCCTTCGAGCGTCCTCCTCATC 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28225 AAAAGGCTTCTACCGCAGAGGGAAGCAGAGGCGGCAGTCGAGTCCTTCGAGCGTCCTCCTCATC 28284

Qy 181 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGCAGCAGTTCTCCTGC 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28285 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGCAGCAGTTCTCCTGC 28344

```
REFERENCE   1  (bases 1 to 29530)
AUTHORS     .
CONSRTM     The SARS epidemiology consortium of Guangdong
TITLE       From independent foci of epidemic outbreak to large genomic
            alteration in late phase viruses: evolution of the SARS-coronavirus
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 29530)
AUTHORS     .
CONSRTM     The SARS epidemiology consortium of Guangdong
TITLE       Direct Submission
JOURNAL     Submitted (19-SEP-2003) Guangdong, China
FEATURES             Location/Qualifiers
source      1  .29530
            /organism="SARS coronavirus HSZ-Bb"
            /mol_type="genomic RNA"
            /isolate="HSZ-Bb"
            /db_xref="taxon:249070"
ORIGIN

Query Match         100.0%;    Score 1136;   DB 14;    Length 29530;
Best Local Similarity  100.0%;    Pred. No. 0;     Indels    0;    Gaps    0;
Matches 1136;   Conservative   0;   Mismatches    0;   Indels    0;   Gaps    0;

Qy      1  AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACCATTGGCAC  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28285  AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACCATTGGCAC  28344

Qy     61  CCGCCAATCCTAATAACAATGCTACCACCGTGCTACAAACTTCCTCAAGGAACAACATTGCC  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28345  CCGCCAATCCTAATAACAATGCTACCACCGTGCTACAAACTTCCTCAAGGAACAACATTGCC  28404

Qy    121  AAAAGGCTTCTACGGCAGAGAGGGAAGCAGAGAGCGGCAGTCAAGCCTCTTCTTCTGCTCCTCATC  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28405  AAAAGGCTTCTACGGCAGAGAGGGAAGCAGAGAGCGGCAGTCAAGCCTCTTCTTCTGCTCCTCATC  28464

Qy    181  ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGCAGCAGTAGGGGAAATTCTCCTGC  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28465  ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGCAGCAGTAGGGGAAATTCTCCTGC  28524

Qy    241  TCGAATGGCTAGCCGAGGTGGTGAAACTGCCCTCGCCCTATTGCTGCTAGACAGATTGAA  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28525  TCGAATGGCTAGCCGAGGTGGTGAAACTGCCCTCGCCCTATTGCTGCTAGACAGATTGAA  28584

Qy    301  CCAGCTTGAGAGCAAAGTTTCTGGTAAAGGCCAACAACAAGGCCAACAACAAGGCCAAACTGTCACTAA  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28585  CCAGCTTGAGAGCAAAGTTTCTGGTAAAGGCCAACAACAAGGCCAACAACAAGGCCAAACTGTCACTAA  28644

Qy    361  GAAATCTGCTCTGAGGCATCTAAAAAGCCTCGCCAAAAACGTACTGCCAAAAACAGTA  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28645  GAAATCTGCTCTGAGGCATCTAAAAAGCCTCGCCAAAAACGTACTGCCAAAAACAGTA  28704

Qy    421  CAACGTCACTCAAGCATTTGGGAGACGTGGTCCAGAACAAACCCAAGGAAATTTCGGGGA  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28705  CAACGTCACTCAAGCATTTGGGAGACGTGGTCCAGAACAAACCCAAGGAAATTTCGGGGA  28764

Qy    481  CCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGCCGCCGCAAATTGCACAATTGC  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28765  CCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGCCGCCGCAAATTGCACAATTGC  28824

Qy    541  TCCAAGTGCCTCTGCAATTCTTGGAATGTCACGCATTGGAGTCACACCTTCGGG  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28825  TCCAAGTGCCTCTGCAATTCTTGGAATGTCACGCATTGGAGTCACACCTTCGGG  28884

Qy    601  AACATGGCTGACTTATCATCGGAGCCATTAAATTGGATGACAAGATCCACAATTCAAAGA  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28885  AACATGGCTGACTTATCATCGGAGCCATTAAATTGGATGACAAGATCCACAATTCAAAGA  28944

Qy    661  CAACGTCATATACTGCTGAACAAGCACATTGACGCATACAAAAACATTCCCACCAACAGAGCC  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28945  CAACGTCATATACTGCTGAACAAGCACATTGACGCATACAAAAACATTCCCACCAACAGAGCC  29004

Qy    721  TAAAAAGGACAAAAAGAAAAGAAAGACTCAGCCTTTGCCGCAGAGACAAAAGA  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  29005  TAAAAAGGACAAAAAGAAAAGAAAAGACTGATGAAGCTCAGCCTTTGCCGCAGAGACAAAAGAA  29064

Qy    781  GCAGCCCACTGTGACTCTTCCTGCGGCTGACATGGACGATGATTCTTCCAGACAACTTCA  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29065  GCAGCCCACTGTGACTCTTCCTGCGGCTGACATGGACGATGATTCTTCCAGACAACTTCA  29124

Qy    841  AAATTCCATGAGTGGAGCTTCGCTGCTGCTTCAACTCAACTCAACACTCATGATGACCAC  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29125  AAATTCCATGAGTGGAGCTTCGCTGCTGCTTCAACTCAACTCAACACTCATGATGACCAC  29184

Qy    901  ACAAGGCAGATGGCTATGTAAACGTTTCGCAATTCCGTTTACGATACATAGTCTACTC  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29185  ACAAGGCAGATGGCTATGTAAACGTTTCGCAATTCCGTTTACGATACATAGTCTACTC  29244

Qy    961  TTGTGCAGAATGAATTCGTAACTAAACAGCACAAGTAGGTTAGTTAACTTTAATCTC  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29245  TTGTGCAGAATGAATTCTCGTAACTAAACAGCACAAGTAGGTTAGTTAACTTTAATCTC  29304

Qy   1021  ACATAGCAATCTTTAATCAATGTGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT  1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29305  ACATAGCAATCTTTAATCAATGTGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT  29364

Qy   1081  TCATCGAGGCCACGCGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC  1136
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29365  TCATCGAGGCCACGCGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC  29420
```

```
RESULT 14
AY338174
LOCUS       AY338174               29573 bp    RNA     linear   VRL 28-JUL-2003
DEFINITION  SARS coronavirus Taiwan TC1, complete genome.
ACCESSION   AY338174
VERSION     AY338174.1  GI:32493129
KEYWORDS    .
SOURCE      SARS coronavirus Taiwan TC1
ORGANISM    SARS coronavirus Taiwan TC1
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 29573)
AUTHORS     Chang,J.-G.C., Lin,T.-H., Chen,C.-M., Lin,C.-S., Chan,W.-L. and
            Shih,M.-C.
TITLE       SARS coronavirus TC1, clinical specimen
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 29573)
AUTHORS     Chang,J.-G.C., Lin,T.-H., Chen,C.-M., Lin,C.-S., Chan,W.-L. and
            Shih,M.-C.
TITLE       Direct Submission
JOURNAL     Submitted (08-JUL-2003) Department of Molecular Medicine, China
            Medical University Hospital, 2, Yuh Der Road, Taichung, Taichung
            404, Taiwan
REFERENCE   3  (bases 1 to 29573)
AUTHORS     Chang,J.-G.C., Lin,T.-H., Chen,C.-M., Lin,C.-S., Chan,W.-L. and
            Shih,M.-C.
TITLE       Direct Submission
JOURNAL     Submitted (28-JUL-2003) Department of Molecular Medicine, China
            Medical University Hospital, 2, Yuh Der Road, Taichung, Taichung
            404, Taiwan
REMARK      Amino acid sequence updated by submitter
FEATURES             Location/Qualifiers
source      1  .29573
            /organism="SARS coronavirus Taiwan TC1"
            /mol_type="genomic RNA"
            /isolate="TC1"
            /db_xref="taxon:237639"
            /note="clinical specimen"
CDS         join(196. .13329,13329. .21416)
            /codon_start=1
            /product="nonstructural polyprotein pp1ab"
            /protein_id="AAQ01594.1"
            /db_xref="GI:3285967"
            /translation="MESIVLGVNEKTHVQLSLPVLQVRDLVLRFGDSVEEALSEARE
            HLKNGTCGLVELEKGVLPQLEQPVVFIKRSDALSTNHGHKVVELVAEMDGIQYGRSGI
            TLGVLVPHVGETPIAYRNVLLRKNGNKGAGGHSYGIDLKSYGIDLGDELGTDPIEDYEQN
```

```
WNTKHGSGALRELTRELNGGAVTRVYDNNFCGPDGYPLDCIKDFLARAGKSMCTLSEQ
LDYIESKRGVYCCRRDHEHEIAWFTERSDKSYEHQTPEIKSAKKFDTFKGECPKFVFP
LNSKVKVIQPRVEKKTEGFMRIRSVYPVASPQECNNMHLSTLMKCNHCDEVSWQTC
DFLKATCEHCGTENLVIEGPTTCGYLPTNAVVKMPCPACADGPEHSVADYHHNSN
IETRLRKGRTRCFGGGCVFAYVGCYNKRAYWVPRASADIGSGHTGITGDNVETLNEDL
LEILSRERVNINIVGDFHLNEVAIILASFSASTSAFIDTIKSLDYKSFKTIVESCGN
YKVTKGQRAVAGNIGQORSVLTPLCGFPSQAAGVIRSIFARTLDAANHSIPDLQRAA
VTILDGISEQSLRLVDAMVYTSDLLTNSVIIMAYVTGGLVQQTSQWLSNLLGTTVEKL
NKALEMCIDQVTIAGAKLRSLNLGEVFLKNGELEALETPVDSFTNGAIVGTPVCVNGL
RPIFEWIEAKLSAGVEFLKRLKGGAPIKGVTFGEDTVWEVQGYKNVRITFELDERVDKVLNE
YCALSPGLLATNNVFRLKGGAPIKGVTFGEDTVWEVQGYKNVRITFELDERVDKVLNE
KCSVYTVESGTEVTEFACVVAEAVVKTLQPVSDLLTNMGIDLLDEWSVATFYLFDDAGE
ENFSSRMYCSFYPDEEEDDAECEEEEIEDTCEHEYGTEDDYQGLPLEFGASAETVR
VEEEEEDWLDDTTEQSEIEPEPTPEEPVNQFTGYLKLTDNVAIKCVDIVKEAQSA
NPMVIVNAANIHLKHGGGVAGALNKATNGAMQKESDDYIKLNGPLTVGGSCLLSGHNL
AKKCLHVVGPNLNAGEDIQLLKAAYENFNSQDILLAPLLSAGIFGAKPLQSLQVCVQT
VRTQVYIAVNDKALYEQVVMDYLDNLKPRVEAPKQEEPPNTEDSKTEEKSVVQKPVDV
PKPIKACIDEVTTLEETKFLTNKLLLFADINGKLYHDSQNLRGEDMSFLEKDAPYM
VGDVITSGDITCVVIPSKKAGGTEMLSRALKKKVPDYEYITTYPGQCAGYTLEERAKT
ALKKCKSAFYVLPSEAPNAKEEILGTVSWNLREMLAHAEFTRKLMPICMDVRAIMATI
QRKYKGIKIQEGIVDYGVARFFYTSKEPVASIITKLNSLNEPLVIMGPLTQTVTHGFNLE
EAARCMRSLKAPAVVSVSSPDAVTTYNGYLTSSSKTSEEHFVETVSLAGSYRDWSYSG
QRTELGVEFLKRGDKIVYHTLESPVEHHLDGEVTLKPSDEGVKVTQELGHEDLMA
NTNLHTQLVDMSMTYGQOFGPTYLDGADVTKIKPHVNHEGKTFPVLPSDDTLRSEAFE
SAINRQVGTMVTGQOFGPTYLDGADVTKIKPHVNHEGKTFPVLPSDDTLRSEAFE
VLFTQTETAHSCNVNRFNVAITRAKIGILCIMSDRDLYDKLQFFTSLEIPRRNVATLQ
YHHTLDESFLGRYMSALNHTKKWKFPQVGGLTSIKWADNNCYLSSVLLALQQLEVKFN
APALQEAYYRARAGDAANFCALILAYSNKTVGELGDVRETMTHLLQHANLESAKRVLN
VCVKHGQCKTTLTGVEAVMMGTSYDNLKTVGVSIPCVCGRDATQYLVQQESVMM
GFKMNYQVNGYPNMFITREEAIRHVRAWIGFDVGSICHATRDAVGTNLPLQLGFSTGVN
SAPPAEYKLQQGTFLCANEYTGNYQCGHYTHITAKETLYRIDGAHLTKMSEYKGPVTD
VFYKEFTSYTTTIKPVSYKLDGVTYTEIEPKLDGYYKKDNAYYTEQPIDLVPTQPLPNA
SFDNFKLTCSNTKFADDLNQMTGFTKPASRELSVTFFPDLNGDVVAIDYRHYSASFKK
GAKLLIKPIVWHINQATTKTTFKPNTWCLRCLWSTRPVDTSNSFEVLAVEDTQGMDNL
ACESQQPTSEEVVENPTIQKEVIECDVKTTEVVGNVILKPSDEGVKVTQELGHEDLMA
AYVENTSITIKKPNELSLALGLKTIATHGIAINSVPWSKILAYVKPFLGQAAITTSN
CAKRLAQRVFNNYMPYVFTLLFQLCTPTKSTNSRIRASLPTTIAKNSVKSVAKLCLDA
GINYVKSPKFSKLFTIAMWLLLSICLGSLICVTAAFGVLLSNFGAPSYCNGVRELYL
NSSNVTMDFCGSFPCSICLSGDSLDSYPALETIQVTISSYKLDLTILGLAAEWVL
AYMLFTKFFYLLGLSAIMQVFGYFASHFISNSWLMWFIISIVQOMAPVSAMVRMYIFF
ASFYYIWKSYVHIMDGCTSSTCMCYKRNRATRVECTTIVNGMKRSFYVYANGGRGFC
KTHNWNCLNCDTFCGTSTFISDEVARDLSLQFKRPINPTDQSSYIVDSVAVKGALHL
YFDKAGQKTYERHPLSHFVNLDNLRANNTKGSLPINVIVFDGKSKCDESASKSASVYY
SQLMCQPILLLDQALVSDVGDSTEVSVKMFDAYVDTTTLNVLAWLYAAVINGDRWFLNRFTTT
LAKGVALDGVLSTFVSAPAGVVDTDVDTKDVIECLKLSHHSDLEVTGDSCNFMLTY
NKVENMTPRDLGACIDCNARHINAQVAKSHNVSLIWNVKDYMSLSEQLRKQIRSAAKK
NNIPFRLTCATTRQVVNVITTKISLKGGKIVSTCFKLMLKATLLCVLAALVCYIVMPV
HTLSIHDGYTNEIIGYKAIQDGVTRDIISTDDCFANKHAGFDAAVRRVAWTLMLNTVSA
PVVAAIITREIGFIVPGLPGTTVLRAINGDFLHFLPRVFSAVGNICYTPSKLIEYSDFA
TSACVLAAECTIFKDAMGKPVPYCYDTNLLEGSISYSELRPDTRYVLMDGSIIQFPNT
YLEGSVRVVTTPDAEYCRHGTCERSEVGICLSTSGRWVLNNEHYRALSGVFCGVDAMN
LIANIFTPLVQPVGALDVSASVVAGGIIAILVTCAAYYFMKFRRVFGEYNHVVAANAL
LFLMSFTILCLVPAYSFLPGVVSVFYLYLTFYFTNDVSFLAHLQWFAMFSPIVPFWIT
AIYVFCISLKHCHWFFNNYLRKRVMFNGVTFSTFEEAALCTFLLNKEMYLKLRSETLL
PLTQYNRYLALYNKYKYFSGALDTTSYREAACCHLAKALNDFSNSGADVLYQPPQTSI
TSAVLQSGFRKMAFPSSGKVEGCMVQVTVCGTTTLNGLWLDDTTVYCPRHVICTAEDMLNP
TLGVLPHVGETPIAYRNVLLRKNGNKGAGGHSYGIDLLKSYDLGDELGTDPIEDYEQN
NYEDLLIRKSNHSFLVQAGNVQLRVIGHSMQNCLLRLKVDTSVKSVKVVIDYGSVFNIDYDC
TFSVLACYNGSPSGVYQCAMRPNHTIKGSFLNGSCGSVGFNIDYDCVSFCVMHHMELP
TGVHAGTDLEGKFYGPFVDRQTAQAAGTDTTTLNVLAWLYAAVINGDRWFLNRFTTT
LNDFNLVAMKNYEPLTQDHVDILGPLSAQTGIAVLDMCAALKELLQNGMNGRTILGS
TILEDEFTPFDVVRQCSGVTFQGKFKKIVKGTHHWMLLTFLTSLLILVQSTQWSLFFF
VVENAFLPFTLGIMAIAACAMLLVKHKHAFLCLFLLPSLATVAYFNMVYMPASWVMRI
MTWLELADTTSLSGYRLKDCVMASSALVLLILMTARTYDDARRWVWTLMNVTIVYKV
YYGNALDQAISMWALVISVTSNYSGVTTTIMELARAIVFVCVEYYPLLFITGNTLQCI
MLVYCFLGYCCCCYFGLCLLNRYRLITLGVYDYLVSTQEFRYMNSQLLPPKSSIDA
FKLNIKLLGIGKPCIKVATVGSHNITGHPDNLRLCEEMLDNRATLQAIASEFSSLPS
YAAYATAQEAYEQAVANGDSVTYNPKKSLVKKSLNVAKSEFDRDAAMQRKLEKMADQAMTQ
MYKQARSEDKRAKVTSAMQTMLFTMLRKLRDNDALNNIINNARDGCVPLNIIPLTTAAK
LMVVVPDYGTYKNTCDGNTTFYASALMEIQQVVDADSKIVQLSEIRMDNSPNLAWPLI
VTIALRAANSAVKLQNNELSPVALRQMSCAAGTTQTACTDDNALAYYNNSKGGRFVLALL
SDHQDLKWARFPKSDGTGTIYTELEPPCRFVTDTPKGPKVKYLYFIKGLNNLNRGMVL
GSLAATVRLQAGNATEVPANSTVLSFCAFAVDPAKAYKDYLASGGOPITNCVKMLCTH
TGTGQAITVTPEANMDQESFGGASCCLYCRCHIDHPNPKGFCDLKGKYVQITPTCAND
PVGFTLRNTVCTVCGMWKGYGCSCDQLREPLMQSADASTFLNRVCGVSAARLTPCGTG
```

```
TSTDVVYRAFDIYNEKVAGFAKFLKTNCCRFQEKDEEGNLLDSYFVVKRHTMSNYQHE
ETIYNLVKDCPAVAVHDFFKFRVDGDMVPHISRQRLTKYTMADLVALVALRHFDEGNCDT
LKEILVLTNDQDLNGNWYDFGPFQVVAPGCADPGPEHSVADYHHNSN
GIVGVLLTLDNQDLNGNWYDFGPFQVVAPGCADPGPEHSVADYHHNSN
DADLAKPLIKWDLLKYDFTEERLCLFDRYFKYWDQTYHPNCINCLDDRCILHCANFNV
LFSTVFPPTSFGPLVRKIFVDGVPFVVSTGYHFRELGVHNQDVNLHSSRLSFKELLV
YAADPAMHAASGNLLLDKRTTCFSVAALTNNVAALTNNVALTNNVALTNNVALTNNVALTN
GSSVELKHFFAQDGNAAISDYDYRYNLPTMCDIROLLFVVEVVDKYFDCYDGGCIN
ANQVIVNNLDKSAGFPFNKWGKARLYDSMSYEDQDALFAYTKRNVIPTITQMNLKYA
ISAKNRARTVAGSVICSTMVTGMQKARVIGTSKFYGGWHMLKTV
YSDVETPHLMGWDYPKCDRAMPNMLRIMASLVLARKHNTCCNLSHRFYRLANECAQVL
SEMVMCGGSLYVKPGGTTSSGDATTAYANSVFNICQAVTANVNALLSTDGNKIADKYVR
NLQHRLYECLYRNRDVHEFVDEFYAYLRKHFSMMILSDDAVVCYNSNYAAQGLVASI
KNFKAVLYYQNNVFMSEAKCWTETDLTKGPHEFCSQHTMLVKQGDDYVYLPYPDPSRI
LGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKHPNQEYADVFHLYLQYIRKLHDEL
TGHMLDMYSVMLTNDNTSRYWEPEFYEAMYTPHTVLQAVGACVLNCILDRCILHCANFNV
RPFLCKKCCCYDHVISTSHKLVLSVNPYVCNAPGCDVTDVTQLYLGGMSYCKSHKPPI
SPPLCANGQVFGLYKNTCVGSDNVTDFNAIATCDWTNAGDYILANTCTERLKLFAAET
LKATEETFKLSYGIATVREVLSDRELHLSEWVGKRPPLNRNVVFTGYRVTNSKVQI
GEYTFEKGDYGDAVVVRGTTTYKLNVGDYFVLTSHTVMPLSAPTLVPQEHYVRITGLY
PTLNISDEFSSNVANYQKVSTLQGPPGTGKSHFAIGLALYFPSARIVTACSH
AAVDALCEKALKYLPIDKCSRIIPARARVECFDVKNSTLQEPFYYSDSPCESHGKQVVSDID
IVVFDEISMATNYDLSVVNARLRAKHYVYIGDPAQLPAPRTLLTKGTLEPEYFNSVCR
LMKTIGPDMFLGTCRRCPAEIVDTVSALVYDNKLKAHKDKSAQCFKMFYKGVITHDVS
SAINRQVGTMVTGQOFGPTYLDGADVTKIKPHVNHEGKTFPVLPSDDTLRSEAFE
VLFTQTETAHSCNVNRFNVAITRAKIGILCIMSDRDLYDKLQFFTSLEIPRRNVATLQ
AENVTGLFKDCSKIITGLHPTQAPTHLSVDIKFKTEGLCVDIPGIPKDMTYRRLISMM
GFKMNYQVNGYPNMFITREEAIRHVRAWIGFDVGSICHATRDAVGTNLPLQLGFSTGVN
LVAVPTGYVDTENNTEFTRVNAKPPPGDQFKHLIPLMYKGLPWNVRIKIVQMLSDTL
KGLSDRVVFVLWAHGFELTSMKYFVKIGPERTCCLCDKRATCFSTSSDTYACWNHSVG
FDYVYNPFMIDVQQWGFTGNLQSNHDQHCQVHGNAHVASCDAIMTRCLAVHECFVKRV
DWSVEYPIIGDELRVNSACRKVQHMVVKSALLADKFPVLHDIGNPKAIKCVPQAEVEW
KFYDAQPCSDKAYKIEELFYSYATHHDKFTDGVCLFWNCNVDRYPANAIVCRFDTRVL
SNLNLPGCDGGSLYVNKHAFHTPAFDKSAFTNLKQLPFFYYSDSPCESHGKQVVSDID
YVPLKSATCITRCNLGAVCRHHANEYRQYLDAYNMMISAGFSLWIYKQFDTYNLWNT
FTRLQSLENVAYNVVNKGHFDGHAGEAPVSIINNAVTKVDGIDVEIFENKTLPVNV
AFELWAKRNIKPVEIKILNNLGVDIAANTVIWDYKREAPAHVSTIGVCTMTDIAKKP
TESSACSSLTVLFDGRVBEGQVDLFRNARNGVLITEGSVKGLTPSKGPAQASVNGVTLIG
ESVKTQFNYFKKVDGIIQQLPETYFTQSRDLEDFKPRSQMETDFLELAMDEFIQRYKL
EGYAFEHIVYGDFSHGQLGGLHIMIGLAKRSQDSPLKLEDFIPMDSTVGNYFITDAQT
GSSKCVCSVIDLLLDDFVEIIKSQDLSVISKVVKVTIDYAEISFMLWCKDGHVETFYP
KLQASAQWPGVAMPNLYKMQRMLLEKCDLQNYGENAVIPKGIMNNVAKYTQLCQYLN
TLTLAVPYNMRVIHFGAGSDKGVAGTYLRQMLIKDLQNAHVNVGKVTRDGFLDPSDTTLIG
DCATVHTANKWDLIISDMYDPRTKHVTKENDSKEGFFTYLCGFIKQKLALGGSIAVKI
TEHSWNADLYKLMGHFSWWTAFVTNVNASSEAFLIGANYLGKPKEQIDGYTMHANYI
FWRNTNPIQLSSYSLFDMSKFPDLKLRGTAVMSLKENQINDMIYSLLEKGRLIIRENNR
VVVSSDILVNN"
196.   .13344
/codon_start=1
/product="nonstructural polyprotein ppla"
/protein_id="AAQ01595.1"
/db_xref="GI:33285968"
/translation="MESLVLGVNEKTHVQLSLPVLQVRDVLVRGFGDSVEEALSEARE
HLKNGTCGLVELEKGVLPQLEQPYVFIKRSDALSTNHGHKVVELVAEMDGIQYGRSGI
TLGVLVPHVGETPIAYRNVLLRKNGNKGAGGHSYGIDLLKSYDLGDELGTDPIEDYEQN
WNTKHGSGALRELTRELNGGAVTRVYDNNFCGPDGYPLDCIKDFLARAGKSMCTLSEQ
LDYIESKRGVYCCRRDHEHEIAWFTERSDKSYEHQTPEIKSAKKFDTFKGECPKFVFP
LNSKVKVIQPRVEKKTEGFMRIRSVYPVASPEQECNNMHLSTLMKCNHCDEVSWQTC
DFLKATCEHCGTENLVIEGPTTCGYLPTNAVVKMPCPACADGPEHSVADYHHNSN
IETRLRKGRTRCFGGGCVFAYVGCYNKRAYWVPRASADIGSGHTGITGDNVETLNEDL
LEILSRERVNINIVGDFHLNEVAIILASFSASTSAFIDTIKSLDYKSFKTIVESCGN
YKVTKGQRAVAGNIGQORSVLTPLCGFPSQAAGVIRSIFARTLDAANHSIPDLQRAA
VTILDGISEQSLRLVDAMVYTSDLLTNSVIIMAYVTGGLVQQTSQWLSNLLGTTVEKL
NKALEMCIDQVTIAGAKLRSLNLGEVFLKNGELEALETPVDSFTNGAIVGTPVCVNGL
RPIFEWIEAKLSAGVEFLKNGELEALETPVDSFTNGAIVGTPVCVNGL
KCSVYTVESGTEVTEFACVVAEAVVKTLQPVSDLLTNMGIDLLDEWSVATFYLFDDAGE
ENFSSRMYCSFYPDEEEDDAECEEEEIDETCEHEYGTEDDYQGLPLEFGASAETVR
VEEEEEDWLDDTTEQSEIEPEPTPEEPVNQFTGYLKLTDNVAIKCVDIVKEAQSA
NPMVIVNAANIHLKHGGGVAGALNKATNGAMQKSDDYIKLNGPLTVGGSCLLSGHNL
AKKCLHVVGPNLNAGEDIQLLKAAYENFNSQDILLAPLLSAGIFGAKPLQSLQVCVQT
VRTQVYIAVNDKALYEQVVMDYLDNLKPRVEAPKQEEPPNTEDSKTEEKSVVQKPVDV
```

CDS

```
         KPKIKACIDEVTTLEEETKFLTNKLLFADINGKLYHDSQNMLRGEDMSFLEKDAPYM
         VGDVITSGDITCVVIPSKKAGGTTEMLSRALKKVPVDEYITTYPGQGCAGYTLEEAKT
         ALKKCKSAFYVLPSEAPNAKEEILGTVSWNLREMLAHAEETRKLMPICMDVRAIMATI
         QRKYKGIKIQEGIVDYGVRFFFYTSKEPVASIITKLNSLNEPLVTMPIGYVTHGFNLE
         EAARCMRSLKAPAVVSVSSPDAVTYNGYLTSSKTSEEHFVEVSLAGSYRDWSYSG
         QRTELGVEFLKRGDKIVHTLESPVEFHLDGEVLSLDKLKSLLSLREVKTIKVFTTVD
         NTNLHTQLVDMSMTYGQOPFGPTYLDGADVTKIKPHVNHEGKTPFVLPSDDTLRSEAFE
         YYHTLDESFLGRYMSALNHTKKKWFPQVGGLTS1KWADNNCYLSSVLLALQQLEVKFN
         APALQEAYRARAGDAANFCALIIAYSNKTVGELGDVRETMTHLLQHANLESAKRVLN
         VVCKHCGQKTTLTGVEAVMYMGTLSYDNLKTGVSIPCVCGRDATQYLVQQSSFVMM
```

```
Query Match       100.0%;  Score 1136;  DB 14;  Length 29573;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 1136;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
Qy      1 AGGCATCGTATGGGTTGCAACTGAGGGAGCTTGAATACACCCAAAGACCACATTGGCAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28437 AGGCATCGTATGGGTTGCAACTGAGGGAGCTTGAATACACCCAAAGACCACATTGGCAC 28496

Qy     61 CCGGCAATCCTAATACAACATGCTGCCACCGTGCTACAACTTCCTCAAGGAACATTGCC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28497 CCGGCAATCCTAATACAACATGCTGCCACCGTGCTACAACTTCCTCAAGGAACATTGCC 28556

Qy    121 AAAAGGCTTCACGGCAGAGGGAAGCAGAGGCGGCAGTCAGCCTCTTCGCTCCTCATC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28557 AAAAGGCTTCACGGCAGAGGGAAGCAGAGGCGGCAGTCAGCCTCTTCGCTCCTCATC 28616

Qy    181 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGCAGCAGTAGGGGAAATTCCTGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28617 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGCAGCAGTAGGGGAAATTCCTGC 28676

Qy    241 TCGAATGGCTAGGCGGAGGTGGTGAAACTGCCTCGCGCTATTGCTGCTAGACAGATTGAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28677 TCGAATGGCTAGGCGGAGGTGGTGAAACTGCCTCGCGCTATTGCTGCTAGACAGATTGAA 28736

Qy    301 CCAGGTTGAGAGCAAAGTTTCTGGTAAAGGCCAACAACAAGGCCAAACTGTCACTAA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28737 CCAGGTTGAGAGCAAAGTTTCTGGTAAAGGCCAACAACAAGGCCAAACTGTCACTAA 28796

Qy    361 GAAATCTGCTGCTGAGGCATCTAAAAAGCCTCGCGCCAAAAACGTACTGCCACAAACAGTA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28797 GAAATCTGCTGCTGAGGCATCTAAAAAGCCTCGCGCCAAAAACGTACTGCCACAAACAGTA 28856

Qy    421 CAACGTTCACTCACTCAAGCAGCATTTGGGAGGACGTGGTCCAGAACAAACCCAAGGAAATTCGGGGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28857 CAACGTTCACTCACTCAAGCAGCATTTGGGAGGACGTGGTCCAGAACAAACCCAAGGAAATTCGGGGA 28916

Qy    481 CCAAGACCTAATCAGACAAGAGAACTGATTACAAAGCGGCCGCAAATTGCACAAATTGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28917 CCAAGACCTAATCAGACAAGAGAACTGATTACAAAGCGGCCGCAAATTGCACAAATTGC 28976

Qy    541 TCCAAGTGCCCTCGCATTCTTCGCACGCATTGGCATGGCAAGTCACACCTTCGGG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28977 TCCAAGTGCCCTCGCATTCTTCGCACGCATTGGCATGGCAAGTCACACCTTCGGG 29036

Qy    601 AACATGGCTGACTTATCATGGAGCCATTAAATTGGATGACAAAGATCCAACAATTCCAAAGA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29037 AACATGGCTGACTTATCATGGAGCCATTAAATTGGATGACAAAGATCCAACAATTCCAAAGA 29096

Qy    661 CAAGCCCACTGTGACTTCTTCCTGCGGCTTGACATGATGATTTCTCCAGACAACTTCA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29097 CAAGCCCACTGTGACTTCTTCCTGCGGCTTGACATGATGATTTCTCCAGACAACTTCA 29156

Qy    721 TAAAAAGGACAAAAAGAAAGACTGATGAAGCTCAGCCTTTGCCGCAGACGAGACAAAGAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29157 TAAAAAGGACAAAAAGAAAGACTGATGAAGCTCAGCCTTTGCCGCAGACGAGACAAAGAA 29216

Qy    781 GCAGCCCACTGTGACTTCTTCCTGCGGCTTGACATGATGATTTCTCCAGACAACTTCA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29217 GCAGCCCACTGTGACTTCTTCCTGCGGCTTGACATGATGATTTCTCCAGACAACTTCA 29276

Qy    841 AAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAGGCATAAAACACTCATGATGACCAC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  29277 AAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAGGCATAAAACACTCATGATGACCAC 29336

Qy    901 ACAAGGCAGATGGGCTATGTGAAACGTTTCCGCAATTCCGTTTACGATACATAGTCTACTC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29337 ACAAGGCAGATGGGCTATGTGAAACGTTTCCGCAATTCCGTTTACGATACATAGTCTACTC 29396

Qy    961 TTGTGCAGAATGAATTCTGTAACTAACAGCACAACAGTAGGTTAGTTAACTTAATCTC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29397 TTGTGCAGAATGAATTCTGTAACTAACAGCACAACAGTAGGTTAGTTAACTTAATCTC 29456

Qy   1021 ACATAGCAATCTTAATCAATGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29457 ACATAGCAATCTTAATCAATGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 29516

Qy   1081 TCATCGAGGCCACGCGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC 1136
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29517 TCATCGAGGCCACGCGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC 29572
```

```
RESULT 15
AY338175
LOCUS       AY338175        29573 bp    RNA        linear   VRL 28-JUL-2003
DEFINITION  SARS coronavirus Taiwan TC2, complete genome.
ACCESSION   AY338175
VERSION     AY338175.1  GI:32493130
KEYWORDS    .
SOURCE      SARS coronavirus Taiwan TC2
  ORGANISM  SARS coronavirus Taiwan TC2
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 29573)
  AUTHORS   Chang,J., Lin,T., Chen,C., Lin,T., Chan,W. and Shih,M.
  TITLE     SARS coronavirus TC2, clinical specimen
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 29573)
  AUTHORS   Chang,J., Lin,T., Chen,C., Lin,T., Chan,W. and Shih,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-2003) Department of Molecular Medicine, China
            Medical University Hospital, 2, Yuh Der Road, Taichung, Taichung
            404, Taiwan
REFERENCE   3  (bases 1 to 29573)
  AUTHORS   Chang,J., Lin,T., Chen,C., Lin,T., Chan,W. and Shih,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JUL-2003) Department of Molecular Medicine, China
            Medical University Hospital, 2, Yuh Der Road, Taichung, Taichung
            404, Taiwan
REMARK      Amino acid sequence updated by submitter
FEATURES             Location/Qualifiers
     source          1..29573
                     /organism="SARS coronavirus Taiwan TC2"
                     /mol_type="genomic RNA"
                     /isolate="TC2"
                     /db_xref="taxon:237640"
                     /note="clinical specimen"
     CDS             join(196..13329,13329..21416)
                     /codon_start=1
                     /product="nonstructural polyprotein pp1ab"
                     /protein_id="AAQ01606.1"
                     /db_xref="GI:33285979"
                     /translation="MESLVLGVNEKTHVQLSLPVLQVRDVLVRGFGDSVEEALSEARE
                     HLKNGTCGLVELEKGVLPQLEQPYVFIKRSDALSTNHGHKVEVLVAEMDGIQYGRSGI
                     TLGVLVPHVGETPIAYRNVLLRKNGNKGAGGHSYGIDLKSYDLGDELGTDPIEDYEQN
                     WNTKHGSGALRELTRELNGGAVTRYVDNNFCGPDGYPLDCIKDFLARAGKSMCTLSEQ
                     LDYIESKRGVYCCRDHEHEIAWFTERSDKSYEHQTPFEIKSAKKFDTFKGECPKVFP
                     LNSKVKVIQPRVEKKTKEGFMGRIRSVYPVASPQECNNMHLSTLMKCNHCDEVSWQTC
                     DFLKATCEHCGTENLVIEGPTTCGYLPTNAVVKMPCPACQDPEIGPEHSVADYHNHSN
                     IETRLRKGGRTRCRFGGCVFAYVGCYNKRAYWVPRASADIGSGHTGITGDNVETLNEDL
                     LEILSRERVNINIVGDFHLNEEVAILASFSASTSAFIDTKLSDLDYKSFKTIVESCGN
                     YKVTKGPVKGAWNIGQQRSVLTPLCGFPSQAAGVIRSIFARTLDAANHSIPDLQRAA
                     VTILDGISEQSLRLVDAMVYTSDLLTNSVIIMAYVTGGLVKGQIQVASDNIKDCVKCFIDVV
                     RPIFEWIEAKLSAGVEFLKDAWEILKFLITGVFDIVKGQIQVASDNIKDCVKCFIDVV
                     NKALEMCIDQVTIAGAKLRSLNLGEVFIAQSKGLYRQCIRGKEQLQLLMPLAPKEVT
                     FLEGDSHDTVLTSEEVLKNGELEALETEPVDSFTNGAIVGTPVCVNGLMLLEIKDKEQ
```

SEMVMCCGGSLYVKPGGTSSGDATTAYANSVFNICQAVTANVNALLSTDGNKIADKYVR
NLQHRLYECLYRNRDVDHEFVDEFYALRKHFSMMLSDDAVVCYNSNYAAQGLVASI
KNPKAVLLYYQNNVFMSEAKCWTEDTLTKGPHEFCSQHTMLVKQGDDYVVLPYPDPSRI
LGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKHPQNEYADVFHLYIQYIRKLHDEL
TGHMLDMYSVMLTNDNTSRYWEPEFYEAMYTPHTVLQAVGACVLCNSQTSLRCGACIR
RPFLCCKCCYDHVISTSHKLVLSVNPYVCNAPGCDVTDVTQLYLGMSYYCKSHKPPI
SFPLCANGAQVDLYKNTCVGSDNVTDFNAIATCDWTNAGDYILANTCERLKLFAAET
LKATEETFKLSYGIATVREVLSDRELHLSWEVGKPRPPLNRNYVFTGYRVTKNSKVQI
GEYTFEBKGDYGDAVVVRGTTYTKLNVGDYFVLTSHTVMPLSAPTLVPQEHYVRITGLY
PTLNIGDEFSSWNARLRAKHYVIGDPAQLPAPRTLLTKGTLEPEYFNSVCR
AAVDALCEKALKYLPIDKCSRIIPARARVECFDKFKVNSTLEQYVFCTVNALPETTAD
IVVFDEISMATNYDLSVNARLRAKHYYIGDPAQLPAPRTLLTKGTLEPEYFNSVCR
LMKTIGPDMFLGTCRRCPAEIVDTVSALVDNKLKAHKDKSAQCFKMPFKGVITHDVS
SAINRPOIGVVREFLTRNPAWRKAVFISPYNSQNAVASKILGLPTQVVDSSQGSEYDY
VIFTQTTETAHSCNVNRFNVAITRAKIGILCIMSDRDLYDKLQFTSLEIPRRNVATLQ
AENVTGLFKDCSKIITGLHPTQAPTHLSVDIKFKTEGLCVDIPGIPKDMTYRRLIISMM
GFKMNYQVNGYPNMFITREEAIRHVRAWIGFDVEGCHATRDAVGTNLPLQLGFSTGVN
LVAVPTGYVDTENNTEFTRVNAKPPEGDQFKHLIPLMYKGLPWNVVRIKIVQMLSDTL
KGLSDRVVFVLWAHGFELTSMKYFVKLGPERTCCLCDKRATCFSTSSDDTYACWNHSVG
FDYVYNPFMIDVQOWGFTGNLQSNHDQHCQVHGNAHVASCDAIMTRCLAVHECFVKRV
DWSVEYPIIGDELRVNSACRKVQHMVVKSALLADKFPVLHDIGNPKAIKCVPQAEVEW
KFYDAQPCSDKAYKIEELFYSYATHHDKFTDGVCLFWNCNVDRYPANAIVCRFDTRVL
SNLNLPGCDGGSLYVNKCAFHTPAFDKSAFTNLKQLPFFYYSDSPCESHGKQVVSDID
YVPLKSATCITRCNLGGAVCRHHANEXRQYLDAYNMMISAGFSLWIYKQFDTYNLWNT
FTRLQSLENVATNVNVRGHFSWWTAFVTNVNASSEAFLIGANYLGKPEQIDGYTMHANYI
APELWAKRNIKPVPEIKILNNLGVDIAANTVIWDYKREAPAHVSTIGVCTMTDIAKKP
TESACSSLITVLFDGRVEGQVDLFRNARNGVLITEGSVGLTPSKGPAQASVNGVTLIG
ESVKTQPNYFKKVDGIIQQLPETYFTQSRDLEDFKPRSQMETDFLELAMDEFIQRYKL
EGYAFEHIVYGDFSHGQLGGLHLMIGLAKRSODSPLKLEDFIPMDSTVKNYFITDAQT
GSSKCVCSVIDLLLDDFVEIIKSQDLSVISKVVKTIDYAEISFMLWCKDGHVETFYP
KLQASQAWQPGVAMPIELYLDKMQRMLLEKCDLQNGENAVIPKGIMMNVAKYTQLCQYLN
TLTLAVPYNMRVIHFGAGSDKGVAPGTAVLRQWLPTGTLLVDSDLNDFVSDADSTLIG
DCATVHTANKWDLIISDMYDPRTKHVTKENDSKEGFFTYLCGFIKQKLALGGSIAVKI
TEHSWNADLYKLMGHFSWWTAFVTNVNASSEAFLIGANYLGKPEQIDGYTMHANYI
FWRNTNPIQLSSYSLFDMSKFPLKLRGTAVMSLKENQINDMIYSLLEKGRLIIRENNR
VVVSSDILVNN"

196. .13344
/codon_start=1
/product="nonstructural polyprotein ppla"
/protein_id="AAQ01607.1"
/db_xref="GI:33285980"
/translation="MESLVLPGGENKTHVQLSLPVLQVRDVLVRGFGDSVEEALSEARE
HLKNGTCGLVELEKGVLPQLEQPYVFIKRSDALSTNHGHKVVELVAEMDGIQYGRSGI
TLGVLVPHVGETPIAYRNVLLRKNGNKGAGGHSYGIDLKSYDLGDELGTDPIEDYEQN
WNTKHGSGALRELTRELNGGAVTRYVDNNFCGPDGYPLDCIKDFLARAGKSMCTLSEQ
LDYIESKRGVYCCRDHEHEIAWFTERSDKSYEHQTPFEIKSAAKFDTFKGECPKFVFP
LNSKVKVIQPRVEKKKTEGFMGRIRSYYPVASPQECNMHLSTLMKCNHCDEVSWQTC
DFLKATCEHCGTENLVIEGPTQGYLPTNAVVHKQPACQDPEIGPEHSVADYHNHSN
IETRLRKGGRTRCFGGCVFAYVGCYNKRAYWVPRASADIGSGHTGITGDNVETLNEDL
LEILSRERVNINIVGDFHLNEEVAIILASFSASTSAFIDTIKSLDYKSFKTIVESCGN
YKVTKGGPVKGAWNIGGQNLTPLCGFPSQAAGVIRSIFARTLDAANHSIPDLQRAA
VTILDGISEQSLRLVDAMVYTSDLLTNSVIIMAYVTGGLVQQTSQWLSNLLGTTVEKL
RPIFEWIEAKLSAGVEFLKDAWEILKFLITGVFDIVKGQICRQKGELKGLYRQCIRGD
NKALEMCIDQVTIAGAKLRSLNLGEVFIAQSKGLYDDSKEDDYIKLNGPLTVGGLQLLMPLKAPKEVT
FLEGDSHDTVLTSEEVVLKNGELEALETPVDSFTNGAIVGTPVCVNGLMLLEIKDKEQ
VCALSPGLLATNNVFRLKGGAPIKGVTFGEDTVWEVQGYKNVRITFELDERVDKVLNE
KCSVYVESGTEVTEFACVVAEAVVKTLQPVSDLLTNMGIDLDEWSVATFYLFDDAGE
ENFSSRMYCSFYPPDEEBEEDDAECEEBEEIDETCEHYGTEDDYQGLPLEFGASAETVR
VEEEBEDWLDDTTEQSEIEPEPTPEEPVNQFTCAND
NPMVIVNAANIHLKHGGGVAGALNKATNGAMQKESDDYIKLNGPLTVGGSCLLSGHNL
AKKCLHVVGPNLNAGEDIQLLKAAYENFNSQDILLAPLLSAGIFGAKPLQSLQVCVQT
QRKYKGIKIQEGIVDYGVRFFFYTSKEPVASIITKLNSLNEPLVTMPIGYVTHGFNLE
EAARCMRSLKAPAVVSVSSPDAVTTVNGYLTTSSSKTSEEHFVETVSLAGSYRDWSYSG
QRTELGVELFKLRGDLVTVYSLALTDKLKJKSLLSLREVKTIKVFTTVD
YHHTLDESFLGRYMSALNHTKKWKFPQVGGLTSIKWADNNCYLSSVILALQQLEVKFN
APALQEAYYRARAGDAANFCALILAYSNKTVGELGDVRETMTHLLQHANLESAKRVLN
VVCKHCGQKTTLTGVEAVMVMGTLSYDNLKTGVSIPCVCGRDATQYLVVQQESSFVMM
SAPPAEYKLQQGTFLCANEYTGNYQCGHYTHITAKETLYRIDGAHLTKMSEYKGPVTD
VFYKETSTTTIKPVSYKLIGPERTCCLCDKRATCFSTSSDDTYACWNHSVG
SFDNFKLTCSNTKFADDLNQMTGFTKPASRELSVTFFDLNGDVVAIDYRHYSASFKK
GAKLLHKPIVWHINQATTKTTFKPNTWCLRCLWSTKPVDTSNSFEVLAVEDTQGMDNL
ACESQQPTSEEVVENPTIQKEVIECDVKTTEVVGNVILKPDSEGVKVTQELGHEDLMA
AYVENTSITIKKPNELSLALGLKTIATHGIAAINSVPWSKILAYVKPFLGQAAITTSN
CAKRLAQRVFNNYMPYVFTILLFQLCFTKSTNSRIRASLPTTIAKNSVKSVAKLCLDA
GINYVKSPKFSKLFTIAMMLLLSICLGSLIVCTAAFGVSLSNFGAPSYCNGVRELYL
NSSNVTTMDFCEGSPFPCSICLSGLDSLDSYPALETIQVTISSYKLDLTILGLAAEWVL
AYMLFKFYYLLGLSAIMQVFGYFASHFISNSWLMWFIISIVQMAPVSAWVRMYIFF
ASFYYIMSVYHIMDGCTSSTCMMCYKRNRATRVECTTIVNGMKRSFYYVANGGRGFC
KTHNWNCLNCDTFCTGSTFISDEVARDLSLQFKRPINPTDQSSYIVFDGKSKCDESASKSASVYY
YFDKAGQKTYERHPLSHFVNLDNLRANNTKGSLPINVIVFDGKSKCDESASKSASVYY
SQLMCQPILLLDQVLVSDVGDSTEVSVKMPDAYVDTFSATFSVPMEKLKALVATAHSE
LAKGVALDGVLSTFVSAARQGVVDTDVDTKDVIECLKLSHHSDLEVTGDSCNNFMLTY
NKVENMTPRDLGACIDCNARHINAQVAKSHNVSLIMNVKDYMSLSEQLRKQIRSAAKK
NNIPFRLTCATTRQVVNVITTKISLKGGKIVSTCFFKLMLKATLLCVLAALVCYIVMPV
HTLSIHDGYTNEIIGYKAIQDGVTRDIISTDDCFANKHAGFDAWFSQRGGSYKNDKSC
PVVAAIITREIGFIVPGLPGTVLRAINGDFLHFLPRVFSAVGNICYTPSKLIEYSDFA
TSACVLAAECTIFKDAMGKPVYCYDTNLLEGSISYSELRPDTRVVLMDGSIIQFPNT
YLEGSVRVTTFDAEYCRHGTCERSEVGICLSTSGRWVLNNEHYRALSGVFCGVDAMN
LIANIFTPLVQPVGALDVSASVVAGGIIAILVTCAAYYFMKFRRVFGEYNHVVAANAL
LFLMSFTILCLVPAYSFLPGVVSVYLYLTFFTNDVSFLAHLQWAMFKRVIQHP
AIYVFCISLKHCWHWFNNYLRKRVMFNGVTFSTFEEAALCTFLLNKEMYLKLRSETLL
PLTQYNRYLALYNKYKYFSGALDTTSYREAACCHLAKALNDFSNSGADVLYQPPQTSI
TSAVLQSGFRKWRKKAFPSGKVEGCMVQVTVCGTTLNGLWLDDTVYCPRHVICTAEDMLNP
NYEDLLIRKSNHSFLVQAGNVQLRVIGHSMQNCLLRLKVDTSNPKTPKYKFVRIQPGQ
TFSVLACYNGSPSGVYQCAMRPNHTIKGSFLNGSCGSVGFNIDYDCVSFCVMHMELP
TGVHAGTDLEGKFYGGPVDRQTAQAAGTDTTITLNVLAWLYAAVINGDRWFLNRFTTT
LNDFNLVAMKYNYEPLTQDHVDILGPLSAQTGIAVLDMCAALKELLQNGMNGRTILGS
TILEDEFTPFDVVRQCSGVTFQGKFKKIVKGTHHWMLLTFLTSLLILVQSTQWSLFFF
VYENAFIPFTLGIMATAACAMLLVKHAFLCLFLLPSLATVAYFNMYMPASWVMRI
MTWLELADTSLSGYRLKDCVMYASALVLLIMTARTVYDDAARRWTLMNVITLVYKV
YYGNALDQAISMWALVISVTSNYSGVVTTIMFLARAIVFVCVEYYPLLFITGNTLQCI
MLVYCPGYCCCCYFGLFCLLNRYFRLTLGVYDYLVSTQEFRYMNSQGLLPPKSSIDA
FKLNIKLLGIGGKPCIKVATVQSKMSDVKCTSVVLLSWMGQAVDINRLCEEMLDNRATLQAISAEFSSLPS
NDILLATTEAFEKMVSLLSVLLSMOGAVDINRLCEEMLDNRATLQAISAEFSSLPS
YAAYATAQEAYEQAVANGDSKRAKVTSAMQTMLFTMLRKLDNDALNNIINNARDGCVPLNIIPLTTAAK
MYKQARSEDKRAKVTSAMQTMLFTMLRKLDNDALNNIINNARDGCVPLNIIPLTTAAK
LMVVVPDYGTYKNTCDGNTFYSAALMEIQQVVDADSKIVQLSEINMDNSPNLAWPLI
VTALRANSAVKLQNNELSPVALRQMSCAAGTTQTACTDNALAYYNNSKGGRFVLALL
SDHQDLKWARFPFKSDGTGTIYTELEEPCRFVTDTPKGPKVKLYLFNRVCGSAARLTPCGTG
PVGFTLRNTVCTVCGMWKGYGCSCDQLREPLMQSADASTFLNRVCGSAARLTPCGTG
TSTDVVYRAFDIYNEKVAGFAKFLKTNCCRFQEBDEGNLLDSYFVVKRHTMSNYQHE
ETIYNLVKDCPAVAVHDFFKFRVDGMVPHISRQRLTKYTMADLVALRHFDEGNCDT
LKEILVTYNCCDDDYFNKKDWYDFVENPDLLRVYANLGERVRQSLLKTVQFCDAMRDA
QRKYKGIKIQEGIVDYGVRFFFYTSKEPVASIITKLNSLNEPLVTMPIGYVTHGFNLE
EAARCMRSLKAPAVVSVSSPDAVTTVNGYLTTSSKTSEEHFVETVSLAGSYRDWSYSG
GIVGVLTLDNQDLNGNWYDFGDFVQVAPGCGVPIVDSYYSLLMPILTLTRALAAESHM
DADLAKPLIKWDLLKDYFTEERLCLRTKFKYFKYDVPTHPNCINCLDDRCILHCANFNV
LFSTVFPPTSFGPLVRKIFVDGVPFVVSTGYHFRELGVVHNQDVNLHSSRLSFKELLV
YAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQTVKPGNFNKDFYDFAVSKGFFKE
GSSVPELKHFFFAQDGNAAISDYGVRNLPTMCDIRQLLFVVEVVDRYFDCYDGGCIN
ANQVIVNNLDKSAGFPFNKWGKARLYDSMSYEDDQDALFAYTKRNVIPTTIQMNLKYA
ISAKNRARTVAGVSICSTMTNRQFHQKLLKSIAATRGATVVIGTSKFYGGWHNMLKTV
YSDVETPHLMGNDYPKCDRAMPNMLRIMASLVLARKHNTCCNLSHRYFLRANECAQVL
CD S

PVVAAIITREIGFIVPGLPGTVLRAINGDFLHFLPRVFSAVGNICYTPSKLIEYSDFA

```
                 VFYKETSYTTIKPVSYKLDGVTYTEIEPKLDGYYKKDNAYYTEQPIDLVPTQPLPNA
                 SPDNFKLTCSNTKFADDLNQMTGFTKPASRELSVTFFPDLNGDVVAIDYRHYSASFKK

Query Match       100.0%;  Score 1136;  DB 14;  Length 29573;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 1136;   Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy       1 AGGCATCGTATGGGTGCAACTGACGGAGCCTTGAATACACCCAAAGACCACATTGGCAC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   28437 AGGCATCGTATGGGTGCAACTGACGGAGCCTTGAATACACCCAAAGACCACATTGGCAC 28496

Qy      61 CCGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGGAACAACATTGCC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   28497 CCGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGGAACAACATTGCC 28556

Qy     121 AAAAGGCTTCTCACGACGAGGAAGCAGAGGCGGCAGTCAAGCCTTCTTCTGCTCCTCATC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   28557 AAAAGGCTTCTCACGACGAGGAAGCAGAGGCGGCAGTCAAGCCTTCTTCTGCTCCTCATC 28616

Qy     181 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGGCAGCAGTAGGGGAAATTCTCCTGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   28617 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGGCAGCAGTAGGGGAAATTCTCCTGC 28676

Qy     241 TCGAATGGCTAGCGGAGGTGTGAAACTGCCCTCGCGCTATTGCTGCTAGACAGATTGAA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   28677 TCGAATGGCTAGCGGAGGTGTGAAACTGCCCTCGCGCTATTGCTGCTAGACAGATTGAA 28736

Qy     301 CCAGCTTGAGCAGCAAAGTTTCTGGTAAAGGCCAACAACAACAAGGCCAAACTGTCACTAA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   28737 CCAGCTTGAGCAGCAAAGTTTCTGGTAAAGGCCAACAACAACAAGGCCAAACTGTCACTAA 28796

Qy     361 GAAATCTGCTGCTGAGGCATCTAAAAAGCTTCGCCAAAAAACGTACTGCCACAAAACAGTA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   28797 GAAATCTGCTGCTGAGGCATCTAAAAAGCTTCGCCAAAAAACGTACTGCCACAAAACAGTA 28856

Qy     421 CAACGTCACTCAAGCATTTGGGAGACGTGGTCCAGAACAACCCAAGGAAATTTCGGGGA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   28857 CAACGTCACTCAAGCATTTGGGAGACGTGGTCCAGAACAACCCAAGGAAATTTCGGGGA 28916

Qy     481 CCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAATTGCACAATTGC 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   28917 CCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAATTGCACAATTTGC 28976

Qy     541 TCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCACACCTTCGGG 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   28977 TCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCACACCTTCGGG 29036

Qy     601 AACATGGCTGACTTATCATGAGCCATTAAATTGGATGACAAAGATCCACAATTCAAAGA 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   29037 AACATGGCTGACTTATCATGAGCCATTAAATTGGATGACAAAGATCCACAATTCAAAGA 29096

Qy     661 CAACGTCATACTGCTGAACAAGCACATTGACGCGCTACAAAACATTCCCACCAACAGAGCC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   29097 CAACGTCATACTGCTGAACAAGCACATTGACGCGCTACAAAACATTCCCACCAACAGAGCC 29156

Qy     721 TAAAAAGGACAAAAAGAAAGAAAAAAAGAAAAAAGAAAGAAGAAGAAGAAGAAGAAAAGAA 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   29157 TAAAAAGGACAAAAAGAAAGAAAAAAAGAAAAAAGAAAGAAGAAGAAGAAGAAGAAAAGAA 29216

Qy     781 GCAGGCCACTGTGACTGTCTTCTTCCCTGCGGCTGACATGATTCTCCAGACAACTTCA 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   29217 GCAGGCCACTGTGACTGTCTTCTTCCCTGCGGCTGACATGATTGATTCTCCAGACAACTTCA 29276

Qy     841 AAATTCCATGAGTGGAGCTTCAACTCAGCCATAAACACTCATGATGACCAC 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   29277 AAATTCCATGAGTGGAGCTTCAACTCAGCCATAAACACTCATGATGACCAC 29336

Qy     901 ACAAGGCAGATGGGCTATGTAAACGTTTCGCAATTCCGTTACGATACATAGTCTACTC 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   29337 ACAAGGCAGATGGGCTATGTAAACGTTTCGCAATTCCGTTACGATACATAGTCTACTC 29396

Qy     961 TTGTGCAGAATGAATTCTCGTAACTAAACAGCACAAACGTAGGTTTAGTTAACTTTAATCTC 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   29397 TTGTGCAGAATGAATTCTCGTAACTAAACAGCACAAACGTAGGTTTAGTTAACTTTAATCTC 29456

Qy    1021 ACATAGCAATCTTTAATCAATGTGTAACATTAGGGAGGACTTGAAAAGAGACTTGAAAAGAGCCACCACATTT 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   29457 ACATAGCAATCTTTAATCAATGTGTAACATTAGGGAGGACTTGAAAAGAGCCACCACATTT 29516

Qy    1081 TCATCGAGGCCACGCGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC 1136
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   29517 TCATCGAGGCCACGCGGAGTACGCGGAGTACAGTGAATAATGCTAGGGAGAGC 29572

Search completed: May 15, 2005, 12:53:28
Job time : 5193.09 secs
```

THIS PAGE LEFT BLANK

```
                    GenCore version 5.1.6
            Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:      May 15, 2005, 10:35:37 ; Search time 713.805 Seconds
                                       (without alignments)
                                       9421.101 Million cell updates/sec

Title:       US-10-764-075-1
Perfect score: 1136
Sequence:    1 aggcatcgtatgggtgcaa........tgaataatgctcagggagagc 1136

Scoring table: IDENTITY_NUC
               Gapop 10_0 , Gapext 1.0

Searched:    4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters:      8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   N_Geneseq_16Dec04:*
          1: geneseqn1980s:*
          2: geneseqn1990s:*
          3: geneseqn2000s:*
          4: geneseqn2001as:*
          5: geneseqn2001bs:*
          6: geneseqn2002as:*
          7: geneseqn2002bs:*
          8: geneseqn2003as:*
          9: geneseqn2003bs:*
         10: geneseqn2003cs:*
         11: geneseqn2003ds:*
         12: geneseqn2004as:*
         13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		1136	100.0	29751	12	ADJ39000	Adj39000 SARS coro
2		48	4.2	2000	8	ADA71938	Ada71938 Rice gene
3		42	3.7	200	8	ABX14055	Abx14055 Mouse POZ
4	c	41.8	3.7	7359	6	ABL33863	Abl33863 Human imm
5		40.4	3.6	463	8	ABX14053	Abx14053 Mouse pox
6		40.2	3.5	2543	6	ABQ43613	Abq43613 Oligonucl
7	c	40.2	3.5	2543	6	ABQ43612	Abq43612 Oligonucl
8		39.2	3.5	9830	6	ABK31241	Abk31241 Signal tr
9		39.2	3.5	9830	6	ABL70196	Abl70196 Chemical
10	c	39.2	3.5	9830	6	AAS61154	Aas61154 Human gen
11	c	39.2	3.5	9830	6	ABN80063	Abn80063 Human che
12	c	39.2	3.4	481	5	ABV57028	Abv57028 Human pro
13		39	3.4	5284	12	ADQ24871	Adq24871 Human sof
14		38.8	3.4	982	12	ADJ74375	Adj74375 Rat cDNA
15		38.6	3.4	2000	8	ADA71938	Ada71938 Rice gene
16	c	38.2	3.3	6001	13	ADS89755	Ads89755 Oligonucl
17	c	38	3.3	8305	6	ABL33568	Abl33568 Human imm
18		37.8	3.3	2748	12	ADP98643	Adp98643 C. albica
19	c	37.6	3.3	236246	12	ADQ97590	Adq97590 Mouse can
20	c	37.4	3.3	1000	9	ADB23206	Adb23206 Environme
21	c	37.4	3.3	2000	6	ABZ16240	Abz16240 Arabidops
22		37.2	3.3	587	6	ABQ21011	Abq21011 Oligonucl
23		37.2	3.3	587	6	ABQ21010	Abq21010 Oligonucl
24	c	37	3.3	551	13	ACN55746	Acn55746 Cotton an
25	c	36.8	3.2	704	5	AAF93473	Aaf93473 cDNA enco
26		36.4	3.2	652	6	ABQ30525	Abq30525 Oligonucl
27		36.4	3.2	652	6	ABQ30524	Abq30524 Oligonucl
28		36.4	3.2	3789	5	AAS78288	Aas78288 DNA encod
29		36.4	3.2	3957	5	AAS78289	Aas78289 DNA encod
30		36.4	3.2	5557	6	ABL33547	Abl33547 Human imm
31		36.4	3.2	7072	6	ABK31470	Abk31470 Signal tr
32		36.4	3.2	7072	6	ABL70565	Abl70565 Chemical
33		36.4	3.2	7072	6	AAS61384	Aas61384 Human gen
34	c	36.4	3.2	11690	9	ACC45236	Acc45236 Human BMC
35		36.2	3.2	353	6	ABL77168	Abl77168 Human ova
36		36.2	3.2	3905	4	AAK67710	Aak67710 Human imm
37		36.2	3.2	3905	4	AAK84071	Aak84071 Human ner
38		36.2	3.2	3905	5	ABA18114	Aba18114 Human ner
39		36.2	3.2	3905	5	ABA21125	Aba21125 Human ner
40		36.2	3.2	3905	12	ADJ12799	Adj12799 DNA fragm
41	c	36.2	3.2	110000	11	ACN43984_0	Acn43984 Mouse gen
42	c	36	3.2	1278	13	ADS89703	Ads89703 Oligonucl
43	c	36	3.2	11209	6	ABN80136	Abn80136 Human che
44		35.8	3.2	11209	8	ABZ36661	Abz36661 Human GEN
45		35.8	3.2	713	4	AAI60951	Aai60951 Human pol

ALIGNMENTS

```
RESULT 1
ADJ39000
ID   ADJ39000 standard; DNA; 29751 BP.
XX
AC   ADJ39000;
XX
DT   06-MAY-2004 (first entry)
XX
DE   SARS coronavirus nucleotide sequence.
XX
KW   small interfering RNA; siRNA; modified ribonucleotide.
KW   viral replication inhibition; hepatitis C virus; HCV; hepatitis C;
KW   antiinflammatory; hepatotropic; virucide; hepatitis A virus;
KW   hepatitis D virus; hepatitis E virus; Ebola virus; influenza virus;
KW   rotavirus; reovirus; retrovirus; poliovirus; human papilloma virus;
KW   metapneumoniavirus; coronavirus; viral infection; gene; ds.
XX
OS   SARS coronavirus.
XX
PN   WO200401647-A1.
XX
PD   05-FEB-2004.
XX
PF   25-JUL-2003; 2003WO-US023104.
XX
PR   26-JUL-2002; 2002US-0398605P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Han J, Seo MY, Houghton M;
XX
DR   WPI; 2004-143862/14.
XX
PT   New RNase resistant small interfering RNA, useful for treating viral
PT   infections, e.g., hepatitis C, influenza virus or coronavirus infection.
XX
PS   Example 10; Fig 3; 74pp; English.
XX
CC   The present invention describes a small interfering RNA (siRNA) which
CC   comprises a modified ribonucleotide, where the siRNA is resistant to
CC   RNase and retains the ability to inhibit viral replication. Also
CC   described: (1) inactivating a virus in a patient; (2) making a modified
CC   siRNA that targets a nucleic acid sequence in a virus; (3) a double-
```

CC stranded RNA molecule of 10-30 nucleotides that inhibits replication of
CC hepatitis C virus (HCV); (4) inducing targeted RNA interference toward
CC HCV in hepatic cells; (5) inhibiting replication of HCV; (6) a vector
CC comprising a DNA segment encoding the RNA molecule; (7) a host cell
CC comprising the vector of (6); (8) inhibiting replication of HCV in cells
CC carrying HCV; (9) treating hepatitis C in a subject; (10) a modified
CC siRNA molecule comprising a double-stranded RNA molecule of 10-30
CC nucleotides in length, which mediates RNA interference toward a target
CC agent or virus and is linked to at least one receptor-binding ligand; and
CC (11) inducing targeted RNA interference in a patient. The modified siRNA
CC molecules have antiinflammatory, hepatotropic and virucide activities.
CC The modified RNA molecules are useful for inactivating virus in mammalian
CC cells. The siRNAs are useful for treating hepatitis C virus, hepatitis A
CC virus, hepatitis D virus, hepatitis E virus, Ebola virus, influenza
CC virus, rotavirus, reovirus, retrovirus, poliovirus, human papilloma
CC virus, metapneumoniavirus or coronavirus infections. The methods of the
CC invention can be used to correct or compensate for cellular physiological
CC abnormalities involved in conferring susceptibility to viral infections
CC in patients and/or alleviate symptoms of a viral infection in patients.
CC The present sequence represents the SARS coronavirus nucleotide sequence,
CC which is used in an example from the present invention.
XX
SQ Sequence 29751 BP; 8478 A; 5941 C; 6188 G; 9144 T; 0 U; 0 Other;

Query Match 100.0%; Score 1136; DB 12; Length 29751;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACCATTGGCAC 60
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28506 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACCATTGGCAC 28565

Qy   61 CCGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGGAACAACATTGCC 120
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28566 CCGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGGAACAACATTGCC 28625

Qy   121 AAAAGGCTTCTACGGCAGGAGCAGAGGCGGCAGTCGAGGCTCTTCGCTCCTCATC 180
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28626 AAAAGGCTTCTACGGCAGGAGCAGAGGCGGCAGTCGAGGCTCTTCGCTCCTCATC 28685

Qy   181 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGGCAGTAGGGGAAATTCTCCTGC 240
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28686 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGGCAGTAGGGGAAATTCTCCTGC 28745

Qy   241 TCGAATGGCTAGCCGGAGGTGGTGAAACTGCCCTCCGCCGTATTGCTAGACAGATTGAA 300
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28746 TCGAATGGCTAGCCGGAGGTGGTGAAACTGCCCTCCGCCGTATTGCTAGACAGATTGAA 28805

Qy   301 CCAGCTTGAGAGAGCAAAGTTTCTGTAAAGGCCAACACACAAGGCCAAACTGTCACTAA 360
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28806 CCAGCTTGAGAGAGCAAAGTTTCTGTAAAGGCCAACACACAAGGCCAAACTGTCACTAA 28865

Qy   361 GAAATCTGCTGCTAAAGGCCATCTAAAAAGCCTGCCAAAAACTACTGCCACAAAACAGTA 420
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28866 GAAATCTGCTGCTAAAGGCCATCTAAAAAGCCTGCCAAAAACTACTGCCACAAAACAGTA 28925

Qy   421 CAACGTCACTCAAGACGTTGGGAGACGTTGGTCCAGAACAAAACCCAAGGAAATTCGGGGA 480
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28926 CAACGTCACTCAAGACGTTGGGAGACGTTGGTCCAGAACAAAACCCAAGGAAATTCGGGGA 28985

Qy   481 CCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGCCGCCAAATTGCACAATTGC 540
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28986 CCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGCCGCCAAATTGCACAATTGC 29045

Qy   541 TCCAAGTGCCTTCGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCACACCTTCGGG 600
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29046 TCCAAGTGCCTTCGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCACACCTTCGGG 29105

Qy   601 AACATGGCTGACTTATCATGGAGCCATTAAATTGGATGACAAAGATCCACAATTCAAAGA 660
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29106 AACATGGCTGACTTATCATGGAGCCATTAAATTGGATGACAAAGATCCACAATTCAAAGA 29165

Qy   661 CAACGTCATCATACTGCTGAACAAGCACATTGACGCACATACAAAAACATTCCCACCAGAGCC 720
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 29166 CAACGTCATCATACTGCTGAACAAGCACATTGACGCACATACAAAAACATTCCCACCAGAGCC 29225

Qy   721 TAAAAAGGACAAAAAGAAAGACTGATGAGCTCAGCCTTGCCGCGAGAGACAAAAGAA 780
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29226 TAAAAAGGACAAAAAGAAAGACTGATGAGCTCAGCCTTGCCGCGAGAGACAAAAGAA 29285

Qy   781 GCAGCCCACTGTGACTCTTCTTCCTCGCGATGATGATTCTCCAGACAACTTCA 840
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29286 GCAGCCCACTGTGACTCTTCTTCCTCGCGATGATGATTCTCCAGACACACTTCA 29345

Qy   841 AAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAGGCATAAACACTCATGATGACCAC 900
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29346 AAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAGGCATAAACACTCATGATGACCAC 29405

Qy   901 ACAAGGCAGATGGGCTATGTAAACGTTTCGCAATTCCGTTACGATACATAGTCTACTC 960
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29406 ACAAGGCAGATGGGCTATGTAAACGTTTCGCAATTCCGTTACGATACATAGTCTACTC 29465

Qy   961 TTGTGCAGAATGAATTCTCGTAACTAAACAGCACAAGTAGGTTTAGTTAACTTTAATCTC 1020
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29466 TTGTGCAGAATGAATTCTCGTAACTAAACAGCACAAGTAGGTTTAGTTAACTTTAATCTC 29525

Qy 1021 ACATAGCAATCTTTAATCAATGTGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 1080
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29526 ACATAGCAATCTTTAATCAATGTGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 29585

Qy 1081 TCATCGAGGCCACGCGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGGAGAGC 1136
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29586 TCATCGAGGCCACGCGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGGAGAGC 29641

RESULT 2
ADA71938
ID ADA71938 standard; DNA; 2000 BP.
XX
AC ADA71938;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 5263.
XX
KW Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.
XX
OS Oryza sativa.
XX
PN WO2003000898-A1.
XX
PD 03-JAN-2003.
XX
PF 22-JUN-2001; 2001WO-IB001105.
XX
PR 22-JUN-2001; 2001WO-IB001105.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
DR WPI; 2003-175290/17.
XX
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
PS Claim 27; SEQ ID NO 5263; 899pp; English.
XX
CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to

CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

Query Match   4.2%; Score 48; DB 8; Length 2000;
Best Local Similarity 8.7%; Pred. No. 0.002;
Matches 72; Conservative 357; Mismatches 397; Indels 0; Gaps 0;

Qy 102 CTCAAGGAACACACATTGCCAAAAGGCTTCTACGCAGAGGAAGCAGAGGCCGCAGTCAAG 161
          | :::|:::|| :|: :: :::::| ::::|| ::::::::|| | : | ||:|:::
Db   12 CSSWRGRRRMYMAGMMSCARMGSSRMSRKMGSMSKYRKCSSCGKCKMTTRRKSKWYSASS 71

Qy 162 CCTCTTCTCGCTCCTCATCACGTAGTCGCGTAATTCAAGAAATTCAACTCCTGGCAGCA 221
          : ::::| :: :: ::   ::::::| ::::::::|:
Db   72 ASGRTCGSKWSSGSYSGCKGMKKKRYKRSKRWRGRGRGRCGMRRSRMRWRMGRYRRCARSGRMAGG 131

Qy 222 GTAGGGGAAATTCTCCTGCTCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTCGCGCTAT 281
          | |::| |||:::|::| ::::|: |:: |:|:::::::|| :|:| :|||
Db 132 SGRMWGGKSRMSYWMCYARGCGSCKRKKSKGGSWGKTCRRGARGGSGWSSGAKYKSGSM 191

Qy 282 TGCTGCTAGACAGATTGAACCAGCTTGAGAGCAAAGTTTCTGGTAAAGGCCAACAACAAC 341
          :|::::|: :::::::||: ::::::| ::|: ::|::: |:::::|:
Db 192 SKRMWMSSCGRSGCGRRSAYSRYYGTSRKYGTYKKMTTYSASRCMRAYMTTSYSWACSSY 251

Qy 342 AAGGCCAAACTGTCACTAAGAAATCTGCTGCTGAGGCATCTAAAAAGCCTCGCCAAAAAC 401
          :: : ::: :::: : ::::::::| : :|:::|
Db 252 TWCRSKRRSMMWKMWRKMRMWRSRSYGWYSWSYKMWCTAYKKSYYSRWCYMYRGGWRGAT 311

Qy 402 GTACTGCCACAAAACAGTACAACGTCACTCAAGCATTTGGGAGACGTGGTCCAGAACAAA 461
          :|::: |: |::|::|: |:| | :::::::|: ::::::::::|:::::||
Db 312 RYWGRGYMSRMAMMYKKMYMWYRGYKGMKGRGWWAGRMMWRSMCRWSKACYYMRWRMWRMTR 371

Qy 462 CCCAGAGAAATTTCGGGACCAAGACCTAATCAGACAGGAACTGATTACAAACATTGGC 521
          ::::|: |::::| : :|: : :::: |: ::| ::| |: :|: |::::::
Db 372 RRRWAKKSSRTSRRKRKKRWCMRKRKYKRMRGYSRMRSCKRARWMKRCRSGRAWKMGCRGC 431

Qy 522 CGCAAATTGCACAATTTGCTCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCA 581
          : || :|:|| |::|:| | :|::: :: |:::|: |: ::|| |
Db 432 MTCRMKSYGMWRWKSWKRMASKYKMWSRMYRWRKKKCSRTTMWGKTRGGMWGTMGRCRYK 491

Qy 582 TGGAAGTCACACCTTCGGGAACATGCTGACTTATCATCGAGCCATTAAATTGGATGACA 641
          :|::| ::::::|: | ::::|:|| |::|||| |::::::::: :: |:::|:|
Db 492 KRSGMKRKCRRRRWGRMYRMWRKRYMSARYTMRYCARKKYSYSAARKARCWYRGKGYYW 551

Qy 642 AAGATCCACAATTCAAAGACAACGTCATACTGCTGAACAAGCACATTGACCGCATACAAAA 701
          || :|::::::|:||| ::| ::|::| | :|::||:::|||:|
Db 552 AGMWMKRYKRMYMYKMMWWYKRKYSKCSWYCKMSSYASCMKSARKAGAKMCKRSKMSAWS 611

Qy 702 CATTCCACCAACAGAGCCTAAAAGGACAAAAGACTGATGAAGCTCAGCCTT 761
          ::::: |::::::::|: :: | ::::|:|:|||:::::|
Db 612 KSMRSSRKCRKCASKRSSAKRYAMMGGMTSGSRMSRWKSYTCYWRKWGSMKSTCTWMYYM 671

Qy 762 TGCCGAGAGACAGAAAGAAGCAGCCCACTGTGACTCTTCTTCCTGCGGCTTGACATGGATG 821
          :|: | :|:::|::::|: |: ::: :: |:|:::: :|:: |::|||||
Db 672 SKYTYAKYGSYWRYRYRAWCMYMWRWYYYRYRSYMTYMAWYTSSTRMAMTGMKYSGRYWT 731

Qy 822 ATTTCTCCAGACAACTTCAAAATTCCATGAGTGAGCTTCTGCTGATTCAACTCAGGCAT 881
          ::::: ::|:::::: :::::::::: ||::::| |::|| | |::
Db 732 SWYKYCKCSWKYRSMWYYWSWWWAKTWKMWRRYATRMWMWYRYSMKWYTWCTMWGYWWY 791

Qy 882 AAACACTCATGATGACCACACAAGGCAGATGGGCTATGTAAACGTT 927
          :: ::::: :: ::::::|: ::::|::: ::::::
Db 792 WWRTYMKMRYMWYKCTKTYWYWSATYWTGTWAAWMAKTKMRMGMT 837

RESULT 3
ABX14055
ID ABX14055 standard; DNA; 200 BP.

XX
AC ABX14055;
XX
DT 20-FEB-2003 (first entry)
XX
DE Mouse POZ domain disruption construct, 3' arm targeting sequence.
XX
KW mouse; ds; POZ; poxvirus and zinc finger domain; transgenic;
KW schizophrenia; POZ disruption; neuroleptic; 129/OlaHsd; C57BL/6;
KW cognitive abnormality; decreased PPI response; prepulse inhibition;
KW weight-type abnormality; decreased body weight; decreased organ weight;
KW decreased organ-to-body weight ratio.
XX
OS Mus sp.
OS Synthetic.
XX
PN WO200279422-A2.
XX
PD 10-OCT-2002.
XX
PF 29-MAR-2002; 2002WO-US009714.
XX
PR 29-MAR-2001; 2001US-0280372P.
PR 24-SEP-2001; 2001US-024590P.
XX
PA (DELT-) DELTAGEN INC.
XX
PI Guenther C;
XX
DR WPI; 2003-067440/06.
XX
PT New transgenic mouse comprising a disruption in a POZ domain gene, useful
PT as a model for disease and for identifying agents that modulate gene
PT expression and function, which are potential agents for treating
PT schizophrenia.
XX
PS Example 1; Fig 3; 68pp; English.
XX
CC The invention relates to a transgenic mouse comprising a disruption in a
CC novel poxvirus and zinc finger (POZ) domain gene, where there is no
CC native expression of the POZ domain gene. The disrupted gene is created
CC in 129/OlaHsd mouse embryonic stem cells. The mice are then bred with
CC C57BL/6 female mice to create the F1 generation. The F2 homozygous
CC mutants are created by crossbreeding the F1 mice. The phenotype of the
CC transgenic mouse compared to wild type includes a cognitive abnormality
CC shown by a decreased prepulse inhibition (PPI) response to a startle
CC response test. The transgenic mouse can also show weight-type
CC abnormalities such as decreased body weight, decreased organ weights
CC (e.g. spleen, liver, kidney, thymus and heart), decreased organ-to-body
CC weight ratios (e.g. spleen, liver, kidney, thymus and heart-to-body
CC weight ratios). The transgenic mouse is useful as a model for disease and
CC for identifying agents that modulate gene expression and function, which
CC are potential agents for treating schizophrenia. The present sequence
CC represents the sequence on the 3' arm of the disruption construct which
CC targets the POZ domain gene
XX
SQ Sequence 200 BP; 66 A; 46 C; 45 G; 43 T; 0 U; 0 Other;

Query Match   3.7%; Score 42; DB 8; Length 200;
Best Local Similarity 66.7%; Pred. No. 0.043;
Matches 60; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 177 CATCACGTAGTGCGCGTAATTCAAGAAATTCAAGACTCCTGGCAGCAGTAGGGGAAATTCTC 236
          ||||||| | |||||||||||||||||||| || ||||||||| |||||||||||||
Db   36 CTTCTACTACCCGAAGGATTTCGAGAAAATCAACTTCCTTGCTTCCTTCCTTCCAAGAAGCAGAGAATTCTTC 95

Qy 237 CTGCTCGAATGGCTAGCGCAGGTGGTCAGTGAAA 266
          |||| ||||||| ||||| ||||||||
Db   96 CAGCTCAAGGGGCTGGCAGAGAGGAGGTGAAA 125

RESULT 4
ABL33863/c

ID   ABL33863 standard; DNA; 7359 BP.
XX
AC   ABL33863;
XX
DT   26-MAR-2002 (first entry)
XX
DE   Human immune system associated gene SEQ ID NO: 1836.
XX
KW   Human; immune system disease; cytosine methylation; antiasthmatic;
KW   antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW   neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW   antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW   antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW   acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW   neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW   dB.
XX
OS   Homo sapiens.
XX
PN   WO200200928-A2.
XX
PD   03-JAN-2002.
XX
PF   02-JUL-2001; 2001WO-EP007537.
XX
PR   30-JUN-2000; 2000DE-01032529.
PR   01-SEP-2000; 2000DE-01043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A, Piepenbrock C, Berlin K;
XX
DR   WPI; 2002-130909/17.
XX
PT   Nucleic acid comprising fragment of chemically modified gene, useful for
PT   diagnosis and treatment of diseases associated with abnormal cytosine
PT   methylation.
XX
PS   Claim 1; SEQ ID NO 1836; 32pp + Sequence Listing; German.
XX
CC   The present invention provides a number of human immune system associated
CC   genes which are modified by the methylation of cytosines. The sequences
CC   can be used in the diagnosis and treatment of immune system disorders,
CC   including eye diseases such as retinopathy, neovascular glaucoma and
CC   macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC   leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC   rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC   diseases. The present sequence is a gene of the invention
XX
SQ   Sequence 7359 BP; 1715 A; 200 C; 1879 G; 3565 T; 0 U; 0 Other;

Query Match   3.7%; Score 41.8; DB 6; Length 7359;
Best Local Similarity   55.0%; Pred. No. 0.28;
Matches   82; Conservative   0; Mismatches   67; Indels   0; Gaps   0;

```
Qy    629 AAATTGGATGACAAAGATCCACAATTCAAAGACAACGTCATACTGCTGAACAAGCACATT 688
                ||||  ||| |||| ||   |||||||||   |||| |  |   |||   |||| |||||
Db   1338 AAACTAAATCAAACAAATAAAAAACTATTCAAACATAACGACAAAAAAAAACTATAAAAAT 1279

Qy    689 GACGCATACAAAAACATTCCCACCACAGAGCCTAAAAAGGACAAAAAGGACAAAAAGACTGAT 748
               || |||| | || ||| ||||  ||||||   |||||| |  |||||| ||  | ||
Db    689 GAGGCATACAAAAACATTCCCACCACAGAGCCTAAAAAGGACAAAAAGGACAAAAAGACTGAT 748

Qy    749 GAAGCTCAGCCTTTGCCGCAGAGACAAAA 777
                 |||||    |||
Db   1218 TAATTTTACCATCAAACGAATTATCAAAA 1190
```

RESULT 5
ABX14053
ID   ABX14053 standard; DNA; 463 BP.
XX
AC   ABX14053;
XX
DT   20-FEB-2003 (first entry)
XX
DE   Mouse poxvirus and zinc finger (POZ) domain gene.
XX
KW   mouse; ds; gene; POZ; poxvirus and zinc finger domain; transgenic;
KW   schizophrenia; POZ disruption; neuroleptic; 129/OlaHsd; C57BL/6;
KW   cognitive abnormality; decreased PPI response; prepulse inhibition;
KW   weight-type abnormality; decreased body weight; decreased organ weight;
KW   decreased organ-to-body weight ratio.
XX
OS   Mus sp.
XX
PN   WO200279422-A2.
XX
PD   10-OCT-2002.
XX
PF   29-MAR-2002; 2002WO-US009714.
XX
PR   29-MAR-2001; 2001US-0280372P.
PR   24-SEP-2001; 2001US-0324590P.
XX
PA   (DELT-) DELTAGEN INC.
XX
PI   Guenther C;
XX
DR   WPI; 2003-067440/06.
XX
PT   New transgenic mouse comprising a disruption in a POZ domain gene, useful
PT   as a model for disease and for identifying agents that modulate gene
PT   expression and function, which are potential agents for treating
PT   schizophrenia.
XX
PS   Example 1; Fig 1; 68pp; English.
XX
CC   The invention relates to a transgenic mouse comprising a disruption in a
CC   novel poxvirus and zinc finger (POZ) domain gene, where there is no
CC   native expression of the POZ domain gene. The disrupted gene is created
CC   in 129/OlaHsd mouse embryonic stem cells. The mice are then bred with
CC   C57BL/6 female mice to create the F1 generation. The F2 homozygous
CC   mutants are created by crossbreeding the F1 mice. The phenotype of the
CC   transgenic mouse compared to wild type includes a cognitive abnormality
CC   shown by a decreased prepulse inhibition (PPI) response to a startle
CC   response test. The transgenic mouse can also show weight-type
CC   abnormalities such as decreased body weight, decreased organ weights
CC   (e.g. spleen, liver, kidney, thymus and heart), decreased organ-to-body
CC   weight ratios (e.g. spleen, liver, kidney, thymus and heart-to-body
CC   weight ratios). The transgenic mouse is useful as a model for disease and
CC   for identifying agents that modulate gene expression and function, which
CC   are potential agents for treating schizophrenia. The present sequence
CC   represents the mouse POZ domain gene
XX
SQ   Sequence 463 BP; 145 A; 102 C; 100 G; 116 T; 0 U; 0 Other;

Query Match   3.6%; Score 40.4; DB 8; Length 463;
Best Local Similarity   65.6%; Pred. No. 0.2;
Matches   59; Conservative   0; Mismatches   31; Indels   0; Gaps   0;

```
Qy    177 CATCACGTAGTCGCGGTAATTCAAAGACAACGTCATACTGCTGAACAAGCAGTAGGGGAAATTCTC 236
                ||||  |  ||||  | || |||||  |||| |  ||  |||  ||||| |||||   |||||||| |
Db    112 CTTCTACTACCGAAGAGTTTCGAGAAAATCAACTTCTTGTCTCAAGCAGCAGAATTCTTC 171

Qy    237 CTGCTCGAATGGCTAGCCGGAGGTGGTGAAA 266
               ||||||   |||| |||| || |||||||||
Db    172 CAGCTCAAGGGGCTGGCAGAGGAGGTGAAA 201
```

RESULT 6
ABQ43613
ID   ABQ43613 standard; DNA; 2543 BP.
XX
AC   ABQ43613;
XX

Column 1 (SEQ ID NO 30204 / RESULT 7)

```
DT   12-JUL-2002  (first entry)
XX
DE   Oligonucleotide for detecting cytosine methylation SEQ ID NO 30204.
XX
KW   Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW   drug; side effect; cancer; central nervous system; cardiovascular;
KW   gastrointestinal; respiratory system; single nucleotide polymorphism;
KW   SNP; cell differentiation; ds.
XX
OS   Homo sapiens.
XX
PN   WO200218632-A2.
XX
PD   07-MAR-2002.
XX
PF   01-SEP-2001; 2001WO-EP010074.
XX
PR   01-SEP-2000; 2000DE-01043826.
PR   05-SEP-2000; 2000DE-01044543.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR   WPI; 2002-371829/40.
XX
PT   Determining the degree of cytosine methylation in genomic DNA, useful for
PT   diagnosis and prognosis, comprises selective hybridization of amplicons
PT   from chemically treated DNA.
XX
PS   Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC   This invention describes a novel method for determining the degree of
CC   methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC   genomic sample of DNA. The sample is treated chemically to convert
CC   cytosine (C) but not methylated C, to uracil, then part of the genomic
CC   DNA that contains the target C is amplified to form a labeled amplicon.
CC   The amplicon is hybridised to two classes, each with at least one member,
CC   of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC   degree of hybridisation to both classes is determined from the label on
CC   the amplicon. From the ratio of labels hybridised to the two classes of
CC   oligomers, the degree of methylation is calculated. The method is used:
CC   (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC   and of a wide range of diseases, e.g. cancer, disorders of the central
CC   nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC   particularly by detecting mutations or single nucleotide polymorphisms
CC   (SNP's); and (ii) for differentiation of cell or tissue types and for
CC   investigating cell differentiation. The method allows the methylation
CC   status of many C residues to be determined simultaneously. ABQ13410-
CC   ABQ54121 represent genomic DNA sequences used to illustrate the method
CC   for determining the degree of cytosine methylation described in the
CC   disclosure of the invention
XX
SQ   Sequence 2543 BP; 1529 A; 616 C; 230 G; 168 T; 0 U; 0 Other;

Query Match   3.5%;  Score 40.2;  DB 6;  Length 2543;
Best Local Similarity  59.0%;  Pred. No. 0.52;
Matches   69;  Conservative  0;  Mismatches  48;  Indels  0;  Gaps  0;

Qy   629 AAATTGGATGACAAAGATCCACAATTCAAAGACAACGTCATACTGCTGAACAAGCACATT 688
          ||| ||||| |||||||  |||||  ||||| ||  |||||  |  |||||||||||||
Db    28 AAAAAAACCAAAAAACGGCACAAAATAAATAAACGCACAAAAAACGCACAAAAAAAAAA 87

Qy   689 GACGCATACAAAAACATTCCCACCAACAGAGCCTAAAAAGGACAAAAAGAAAAAGAGACT 745
          ||| |||| |  || || || ||| ||  |  ||||||| || |||| |||| || |
Db    88 AACGCCACGAAAAACACCACCGAAAAAAAACGGCCAAAAAAAACCACTAAAAAAAAAACT 144

RESULT 7
ABQ43612/c
ID   ABQ43612 standard; DNA; 2543 BP.
XX
AC   ABQ43612;
```

Column 2 (SEQ ID NO 30203 / RESULT 8)

```
DT   12-JUL-2002  (first entry)
XX
DE   Oligonucleotide for detecting cytosine methylation SEQ ID NO 30203.
XX
KW   Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW   drug; side effect; cancer; central nervous system; cardiovascular;
KW   gastrointestinal; respiratory system; single nucleotide polymorphism;
KW   SNP; cell differentiation; ds.
XX
OS   Homo sapiens.
XX
PN   WO200218632-A2.
XX
PD   07-MAR-2002.
XX
PF   01-SEP-2001; 2001WO-EP010074.
XX
PR   01-SEP-2000; 2000DE-01043826.
PR   05-SEP-2000; 2000DE-01044543.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR   WPI; 2002-371829/40.
XX
PT   Determining the degree of cytosine methylation in genomic DNA, useful for
PT   diagnosis and prognosis, comprises selective hybridization of amplicons
PT   from chemically treated DNA.
XX
PS   Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC   This invention describes a novel method for determining the degree of
CC   methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC   genomic sample of DNA. The sample is treated chemically to convert
CC   cytosine (C) but not methylated C, to uracil, then part of the genomic
CC   DNA that contains the target C is amplified to form a labeled amplicon.
CC   The amplicon is hybridised to two classes, each with at least one member,
CC   of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC   degree of hybridisation to both classes is determined from the label on
CC   the amplicon. From the ratio of labels hybridised to the two classes of
CC   oligomers, the degree of methylation is calculated. The method is used:
CC   (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC   and of a wide range of diseases, e.g. cancer, disorders of the central
CC   nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC   particularly by detecting mutations or single nucleotide polymorphisms
CC   (SNP's); and (ii) for differentiation of cell or tissue types and for
CC   investigating cell differentiation. The method allows the methylation
CC   status of many C residues to be determined simultaneously. ABQ13410-
CC   ABQ54121 represent genomic DNA sequences used to illustrate the method
CC   for determining the degree of cytosine methylation described in the
CC   disclosure of the invention
XX
SQ   Sequence 2543 BP; 168 A; 230 C; 616 G; 1529 T; 0 U; 0 Other;

Query Match   3.5%;  Score 40.2;  DB 6;  Length 2543;
Best Local Similarity  59.0%;  Pred. No. 0.52;
Matches   69;  Conservative  0;  Mismatches  48;  Indels  0;  Gaps  0;

Qy   629 AAATTGGATGACAAAGATCCACAATTCAAAGACAACGTCATACTGCTGAACAAGCACATT 688
          ||| ||||| |||||||  |||||  ||||| ||  |||||  |  |||||||||||||
Db  2516 AAAAAAACCAAAAAACGGCACAAAATAAATAAACGCACAAAAAACGCACAAAAAAAAAA 2457

Qy   689 GACGCATACAAAAACATTCCCACCAACAGAGCCTAAAAAGGACAAAAAGAAAAAGAGACT 745
          ||| |||| |  || || || ||| ||  |  ||||||| || |||| |||| || |
Db  2456 AACGCCACGAAAAACACCACCGAAAAAAAACGGCCAAAAAAAACCACTAAAAAAAAAACT 2400

RESULT 8
ABK31241/c
ID   ABK31241 standard; DNA; 9830 BP.
XX
```

AC   ABK31241;
XX
DT   23-APR-2002 (first entry)
XX
DE   Signal transduction associated gene modified complementary DNA #42.
XX
KW   Human; signal transduction associated gene; cytosine methylation state;
KW   CpG island; signal transduction associated disease; solid tumour; cancer;
KW   antitumour; cytostatic; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO20200926-A2.
XX
PD   03-JAN-2002.
XX
PF   29-JUN-2001; 2001WO-EP007472.
XX
PR   30-JUN-2000; 2000DE-01032529.
PR   01-SEP-2000; 2000DE-01043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A, Piepenbrock C, Berlin K;
XX
DR   WPI; 2002-147896/19.
XX
PT   Oligonucleotide for diagnosis and therapy of diseases associated with
PT   signal transduction e.g. cancer, comprises chemically modified genomic
PT   sequences of genes associated with signal transduction.
XX
PS   Claim 1; SEQ ID NO 84; 24pp; English.
XX
CC   The present invention relates to chemically modified DNA sequences of
CC   signal transduction associated genes. The DNA sequences are chemically
CC   modified using a solution of bisulphite, hydrogen sulphite or disulphite.
CC   Also disclosed are oligonucleotides and/or PNA oligomers for detecting
CC   the cytosine methylation state (CpG islands) of these genes, and a method
CC   for the diagnosis and/or therapy of genetic and epigenetic parameters of
CC   genes associated with signal transduction. The genomic DNA can be
CC   obtained from cells or cellular components which contain DNA, e.g., cell
CC   lines, biopsies, blood, sputum, stool, urine, cerebral-spinal fluid,
CC   tissue embedded in paraffin such as tissue from eyes, intestine, kidney,
CC   brain, heart, prostate, lung, breast or liver, histologic object slides,
CC   and all their possible combinations. The sequences of the invention are
CC   useful for the diagnosis and therapy of diseases associated with signal
CC   transduction e.g. solid tumours and cancer. ABK31158-ABK31545 represent
CC   chemically pretreated genomic DNA sequences of different genes associated
CC   with signal transduction, or their complementary sequences. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from the
CC   European Patent Office
XX
SQ   Sequence 9830 BP; 2285 A; 267 C; 2773 G; 4505 T; 0 U; 0 Other;

Query Match   3.5%; Score 39.2; DB 6; Length 9830;
Best Local Similarity   52.4%; Pred. No. 2;
Matches   86; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   639 ACAAAGATCCACAATTCAAAGACAACGTCATACTGCTGAACAAGCACATTGACGCATACA 698
              ||| |||| || ||||||||| | || ||| || ||| |||| | |
Db   6121 ACACCAAAACGAAAAACCAAAATATCCACATCCGACAAAAAAAACCCACGTACCTAAATA 6062

Qy   699 AAACATTCCCACCAACAGAGCCTAAAAAGGACAAAAAGAAAAAGACTGATGAAGCTCAGC 758
              | ||| ||| || ||| || |||||||| ||||||| |||| || | |
Db   6061 ACAAATCCCGAACATCAAAACGAAAATAAAATCAAAAAAAAATAAAATAAACAAAAAAAT 6002

Qy   759 CTTTGCCGCAGAGACAAAAGAAGCAGCCCACTGTGACTCTTCTT 802
              || | ||| |||||||| ||| ||| | ||| ||||
Db   6001 CTAAAACCTCCCAAAAAATCAAACCTACCACAATTACTCTTCCT 5958

RESULT 9
ABL70196/c
ID   ABL70196 standard; DNA; 9830 BP.
XX
AC   ABL70196;
XX
DT   01-JUL-2002 (first entry)
XX
DE   Chemically treated cell signalling DNA sequence complementary to#43.
XX
KW   Cell signalling; cytosine methylation; cell signalling disease; cancer;
KW   tumour; cytostatic; ds.
XX
OS   Unidentified.
XX
PN   WO20020807-A2.
XX
PD   10-JAN-2002.
XX
PF   29-JUN-2001; 2001WO-EP007471.
XX
PR   30-JUN-2000; 2000DE-01032529.
PR   01-SEP-2000; 2000DE-01043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A, Piepenbrock C, Berlin K;
XX
DR   WPI; 2002-154758/20.
XX
PT   Nucleic acid, useful for diagnosis and therapy of diseases associated
PT   with cell signaling e.g. cancer, comprises chemically modified genomic
PT   sequences of genes associated with cell signaling.
XX
PS   Claim 1; SEQ ID NO 86; 24pp + Sequence Listing; English.
XX
CC   The invention relates to a nucleic acid comprising a sequence of at least
CC   18 bases of a segment of chemically pretreated DNA of genes associated
CC   with cell signalling. The activity of the modified sequences of the
CC   invention may be described as cytostatic. The object of the invention is
CC   to provide the chemically modified DNA of genes associated with cell
CC   signalling, as well as oligonucleotides and/or PNA-oligomers for
CC   detecting cytosine methylations, as well as a method which is
CC   particularly suitable for the diagnosis and/or therapy of genetic and
CC   epigenetic parameters of genes associated with cell signalling. The
CC   chemically modified DNA provided by the invention is useful for diagnosis
CC   and therapy of diseases such as solid tumours and cancer. The sequences
CC   given in records ABL70111-ABL70626 represent chemically pre-treated
CC   genomic DNA's of genes associated with cell signalling. Note: The
CC   sequence data for this patent is not represented in the printed
CC   specification, but is based on sequence information supplied by the
CC   European Patent Office
XX
SQ   Sequence 9830 BP; 2285 A; 267 C; 2773 G; 4505 T; 0 U; 0 Other;

Query Match   3.5%; Score 39.2; DB 6; Length 9830;
Best Local Similarity   52.4%; Pred. No. 2;
Matches   86; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   639 ACAAAGATCCACAATTCAAAGACAACGTCATACTGCTGAACAAGCACATTGACGCATACA 698
              ||| |||| || ||||||||| | || ||| || ||| |||| | |
Db   6121 ACACCAAAACGAAAAACCAAAATATCCACATCCGACAAAAAAAACCCACGTACCTAAATA 6062

Qy   699 AAACATTCCCACCAACAGAGCCTAAAAAGGACAAAAAGAAAAAGACTGATGAAGCTCAGC 758
              | ||| ||| || ||| || |||||||| ||||||| |||| || | |
Db   6061 ACAAATCCCGAACATCAAAACGAAAATAAAATCAAAAAAAAATAAAATAAACAAAAAAAT 6002

Qy   759 CTTTGCCGCAGAGACAAAAGAAGCAGCCCACTGTGACTCTTCTT 802
              || | ||| |||||||| ||| ||| | ||| ||||
Db   6001 CTAAAACCTCCCAAAAAATCAAACCTACCACAATTACTCTTCCT 5958

RESULT 10

```
AAS61154/c
ID   AAS61154 standard; DNA; 9830 BP.
XX
AC   AAS61154;
XX
DT   29-JAN-2002  (first entry)
XX
DE   Human gene regulation-associated gene oligonucleotide #109.
XX
KW   Human; Gene regulation-associated gene; severe combined immunodeficiency;
KW   cardiac damage; inflammatory response; Haemophilia; Werner syndrome;
KW   asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;
KW   renal disease; Preeclampsia; cardiac allograft vascular disease;
KW   colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;
KW   immunostimulant; cardiant; antiinflammatory; coagulant; antiasthmatic;
KW   nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO200177375-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-EP003968.
XX
PR   06-APR-2000; 2000DE-01019058.
PR   07-APR-2000; 2000DE-01019173.
PR   30-JUN-2000; 2000DE-01032529.
PR   01-SEP-2000; 2000DE-01043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A, Piepenbrock C, Berlin K;
XX
DR   WPI; 2002-017470/02.
XX
PT   New nucleic acid sequences from chemically modified genes associated with
PT   gene regulation, useful for analyzing cytosine methylations for diagnosis
PT   and therapy of diseases e.g. severe combined immunodeficiency disease.
XX
PS   Claim 1; SEQ ID NO 112; 26pp; English.
XX
CC   The invention relates to 224 nucleic acid sequences comprising at least
CC   18 bases of a chemically pretreated gene associated with gene regulation
CC   and in therapy of such diseases, by enabling analysis of the cytosine
CC   methylation patterns of such genes, kits are provided. They are
CC   especially useful in diagnosis and therapy of e.g. severe combined
CC   immunodeficiency disease, cardiac disorders, haemophilia, solid tumours
CC   and cancer, Werner syndrome, asthma, HDR syndrome, Saethre-Chotzen
CC   syndrome, renal disease, preeclampsia, graft versus-host disease. The
CC   present sequence is a sequence included in the sequence data for this
CC   specification and is associated with the human gene regulation-associated
CC   genes. Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 9830 BP; 2285 A; 267 C; 2773 G; 4505 T; 0 U; 0 Other;

Query Match             3.5%;  Score 39.2;  DB 6;  Length 9830;
Best Local Similarity  52.4%;  Pred. No. 2;
Matches   86;  Conservative   0;  Mismatches  78;  Indels   0;  Gaps   0;

Qy        639 ACAAGATCCACAATTCAAAGACACGTCATATCTGCTGACAAGCACATTGACGCATACA 698
              ||| ||||||||| ||||| |||||||  || ||||||||  |||||||||||||||
Db       6121 ACACCAAAACGAAAAACCAAAATATCCACATCCGACAAAAAAAACCCACGTACCTAAATA 6062

Qy        699 AAACATTCCCACCACACAGAGCCTAAAAGGACAAAAAGGACAAAAAGAAAAAGAAAGACTGATGAAGCTCAGC 758
```

```
Db         6061 ACAAATCCCGACCATCAAACCAAAATAAAATCAAAAAATAAACAAAAAAAT 6002
                | ||||||||  ||||    |||||  ||  |||| ||||  ||| |||||| ||||
Qy          759 CTTTGCCGCAGAGACAAAAGAAGAGCAGCGCCCACTGTGACTCTTCTT 802
                   ||||  |||||||   |    |||||   ||||||  |||||||||
Db         6001 CTAAAACTCCCAAAAAATCAACCTACCACATTACTCTTCCT 5958
```

```
RESULT 11
ABN80063/c
ID   ABN80063 standard; DNA; 9830 BP.
XX
AC   ABN80063;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human chemically modified disease associated gene SEQ ID NO 80.
XX
KW   Human; development; homeobox gene; HOX; diabetes; cancer; apoptosis;
KW   heart disease; epilepsy; histone deacetylation; muscular dystrophy;
KW   dwarfism; single nucleotide polymorphism; SNP; cytosine methylation;
KW   antidiabetic; cytostatic; anticonvulsant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO20020927-A2.
XX
PD   03-JAN-2002.
XX
PF   02-JUL-2001; 2001WO-EP007536.
XX
PR   30-JUN-2000; 2000DE-01032529.
PR   01-SEP-2000; 2000DE-01043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A, Piepenbrock C, Berlin K;
XX
DR   WPI; 2002-130908/17.
XX
PT   Novel nucleic acid useful for diagnosis and therapy of diseases
PT   associated with development genes such as diabetes, comprises a sequence
PT   of a segment of chemically pretreated DNA of genes associated with
PT   development.
XX
PS   Claim 1; SEQ ID NO 80; 27pp; English.
XX
CC   The invention relates to a nucleic acid (I) comprising a sequence at
CC   least 18 bases in length of a segment of chemically pretreated DNA (II)
CC   of genes associated with development selected from 87 genes listed in the
CC   specification such as ACCPN, ADFN, or AFD1 and comprising one of 350
CC   sequences (ABN79984-ABN80333) or their complements. The invention is
CC   useful for the diagnosis or therapy of diseases associated with
CC   development genes, in particular disease related to homeobox containing
CC   genes (HOX), like diabetes, cancer, apoptosis related diseases, syndromes
CC   associated with congenital heart disease, epilepsy, diseases related to
CC   histone deacetylation, Currarino syndrome, diseases related with the
CC   development of the brain and limb girdle muscular dystrophy and dwarfism.
CC   Oligomers specific to each of the genes are useful for detecting the
CC   methylation state of all CpG dinucleotides within the 350 sequences or
CC   (II) and their complementary sequences, as primer oligonucleotides for
CC   the amplification of the 350 sequences, (II) and/or their complements and
CC   as oligomer probes for detecting the cytosine methylation state and/or
CC   single nucleotide polymorphisms (SNPs). Note: The sequence data for this
CC   patent did not form part of the printed specification but is based on
CC   sequence information supplied to Derwent by the European Patent Office
XX
SQ   Sequence 9830 BP; 2285 A; 267 C; 2773 G; 4505 T; 0 U; 0 Other;

Query Match             3.5%;  Score 39.2;  DB 6;  Length 9830;
Best Local Similarity  52.4%;  Pred. No. 2;
Matches   86;  Conservative   0;  Mismatches  78;  Indels   0;  Gaps   0;
```

Qy   589 CACACCTTCGGGACATGGCTGACTTATCATGGAGCCATTAAATTGGATGACAAAGATCC 648
          ||| |||| || ||| || || | ||| ||| |||| || |||||||||||
Db   7 CCCCCTCCGGGTACTTCTAGNATTAATTAACGGCGGGATTTAAAAGAAAAAAAAAAC   66

Qy   649 ACAATTCAAAGACAACGTCATACTGCTGAACAAGCACATTGACGCATACAAAACATTCCC 708
          |||||||| || |||
Db   67 AAAAAAAACANAAACANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA   126

Qy   709 ACCAACAGAGCCTAAAAAGGACAAAAAGAAAAAGACTGATGAAGCTCAGCCTTTGCCGCA 768
          | ||| | ||| ||||
Db   127 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTTTTTAAA   186

Qy   769 GAGACAAAAGAAGCAGCCCAC 789
          | || || ||| |||||
Db   187 AAAAAAAAAAAGGCCCCCCC   207

RESULT 13
ADQ24871
ID   ADQ24871 standard; DNA; 5284 BP.
XX
AC   ADQ24871;
XX
DT   26-AUG-2004 (first entry)
XX
DE   Human soft tissue sarcoma-upregulated DNA - SEQ ID 7691.
XX
KW   soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO200048938-A2.
XX
PD   10-JUN-2004.
XX
PF   26-NOV-2003; 2003WO-US038193.
PR   26-NOV-2002; 2002US-0429739P.
XX
PA   (PROT-) PROTEIN DESIGN LABS INC.
XX
PI   Aziz N, Ginsburg WM, Zlotnik A;
XX
DR   WPI; 2004-441208/41.
XX
PT   Early detection of soft tissue sarcoma comprises determining expression
PT   of a gene in a first soft tissue sample and a normal soft tissue sample
PT   and comparing the gene expression, also useful in treating soft tissue
PT   sarcoma.
XX
PS   Example 2; SEQ ID NO 7691; 210pp; English.
XX
CC   The invention relates to a novel method for detecting soft tissue sarcoma
CC   which comprises obtaining a first soft tissue sample from an individual
CC   and a normal soft tissue sample from the same or different individual,
CC   determining the expression of a gene in both samples and comparing the
CC   expression of the gene in both soft tissue samples, where a higher level
CC   of protein expression in the first soft tissue sample indicates the
CC   presence of soft tissue sarcoma. The method of the invention has
CC   cytostatic applications and may be useful for detecting soft tissue
CC   sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC   acid sequences may be useful in diagnostic and screening applications.
CC   The current sequence is that of a human soft tissue sarcoma-upregulated
CC   DNA of the invention. The current sequence is not shown within the
CC   specification per se but was submitted in CD format by the inventor.
XX
SQ   Sequence 5284 BP; 1397 A; 1242 C; 1236 G; 1377 T; 0 U; 32 Other;

Query Match   3.4%; Score 39; DB 12; Pred. No. 1.7; Length 5284;
Best Local Similarity   54.4%;
Matches   98; Conservative   0; Mismatches 81; Indels 1; Gaps 1;

Qy   639 ACAAAGATCCACAATTCAAAGACAACGTCATACTGCTGAACAAGCACATTGACGCATACA 698
          ||| || |||| ||| || ||||||| |
Db 6121 ACACCAAAACGAAAAACCAAAAATATCCACATCCGACAAAAAAACCACGTACCTAAATA 6062

Qy   699 AAACATTCCCACCAACAGAGCCTAAAAAGGACAAAAAGAAAAAGACTGATGAAGCTCAGC 758
          || ||| || ||| | ||||||
Db 6061 ACAAATCCGAACATCAAAAACGAAAAATAAAATCAAAAAAAAAAAAATAAACAAAAAAAT 6002

Qy   759 CTTTGCCGCAGAGACAAAAGAAGCAGCCCACTGTGACTCTTCTT 802
          |||||| | ||||| | ||||| |||
Db 6001 CTAAAACCTTCCCAAAAAATCAAACCTACCACACAATTACTCTTCCT 5958

RESULT 12
ABV57028
ID   ABV57028 standard; cDNA; 481 BP.
XX
AC   ABV57028;
XX
DT   17-SEP-2002 (first entry)
XX
DE   Human prostate expression marker cDNA 57019.
XX
KW   Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW   pharmacogenomic marker; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200160860-A2.
XX
PD   23-AUG-2001.
XX
PF   20-FEB-2001; 2001WO-US005171.
XX
PR   17-FEB-2000; 2000US-0183319P.
PR   16-MAR-2000; 2000US-0189862P.
PR   25-MAY-2000; 2000US-0207454P.
PR   09-JUN-2000; 2000US-0211314P.
PR   18-JUL-2000; 2000US-0219007P.
PR   13-DEC-2000; 2000US-0255281P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R, Endege WO, Monahan JE;
XX
DR   WPI; 2001-662795/76.
XX
PT   Novel isolated nucleic acid molecule associated with cancerous state of
PT   prostate cells and correlating with presence of prostate cancer, useful
PT   for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS   Claim 1; Page 10980; 11750pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule (I) comprising
CC   a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC   specification or its complement. (I) is useful for: (a) assessing whether
CC   a patient is afflicted with prostate cancer; (b) monitoring the
CC   progression of prostate cancer in a patient; (c) assessing the efficacy
CC   of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC   the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC   (e) selecting a composition for inhibiting prostate cancer in a patient;
CC   (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC   determining whether prostate cancer has metastasized in a patient; (h)
CC   assessing the aggressiveness or indolence of prostate cancer in a patient
CC   ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ   Sequence 481 BP; 204 A; 82 C; 48 G; 145 T; 0 U; 2 Other;

Query Match   3.4%; Score 39; DB 5; Pred. No. 0.54; Length 481;
Best Local Similarity   49.3%;
Matches   99; Conservative   0; Mismatches 102; Indels 0; Gaps 0;

```
Qy      665  GTCATACTGCTGAACAAGCACATTGACGCATACAAAACATTCCCACCACAGAG-CCTAA  723
             ||| ||| ||| |||  ||||  ||||            |||||||| |||
Db     5091  GTGACACTTTGGAGCAAAACCTTGCAGTGGTAAATAAAAAATTCCAACAGGGAAAAA   5150

Qy      724  AAAGGACAAAAAGAGAAAAGACTGATGAAGCTCAGCCTTTGCCGCAGAGACAAAAGAAGCA  783
             ||| |||||||||||||||||||||||||||||||||||||||||||| ||||
Db     5151  AAAAAAAAAAAAAAAAAAAAAAACGTTTTACCTAAATATTTGGCATATAAAATGCAGAAGAA   5210

Qy      784  GCCCACTGTGACTCCTTCCTCGCGCTGACATGGATGATTCTCCAGACAACTTCAAAA  843
             ||||| ||||||  ||||||||
Db     5211  ATCATCTGATAAGCACTTATGAGCCNCTGCCATGTTTAATAAACTTGACAACTACAAA  5270

RESULT 14
ADJ74375
ID   ADJ74375 standard; cDNA; 982 BP.
XX
AC   ADJ74375;
XX
DT   06-MAY-2004 (first entry)
XX
DE   Rat cDNA sequence required for viral infection SeqID 289.
XX
KW   ss; viral infection; tumour suppression; bacterial; parasitic growth;
KW   gene trap; serum survival factor; cytostatic; virucidal; antibacterial;
KW   antiparasitic; rat.
XX
OS   Rattus sp.
XX
PN   WO2004010925-A2.
XX
PD   05-FEB-2004.
XX
PF   02-MAY-2003; 2003WO-US013743.
XX
PR   02-MAY-2002; 2002US-0377136P.
XX
PA   (UVVA-) UNIV VANDERBILT.
XX
PI   Rubin DH;
XX
DR   WPI; 2004-143721/14.
XX
PT   Novel gene useful for preparing a composition for reducing or inhibiting
PT   viral infection or for suppressing a malignant phenotype in a cell.
XX
PS   Claim 1; SEQ ID NO 289; 662pp; English.
XX
CC   This invention relates to novel mammalian genes involved in viral
CC   infection and tumour suppression. Specifically, it refers to methods for
CC   identifying cellular genes that are required for viral, bacterial or
CC   parasitic growth, as well as genes used for tumour progression, but that
CC   are not essential for cell survival. The present invention describes a
CC   'gene trap' method and screening/ selection process that can isolate
CC   those polynucleotides that are associated with a specific process of
CC   interest. Furthermore, this method relies on the core discovery that
CC   virally infected cells become dependent on a serum survival factor that
CC   does not affect non-infected cells. Accordingly, identification of these
CC   factors and the appropriate inhibitors thereof, provides a means to treat
CC   and/ or prevent viral, bacterial or parasitic growth and infection, as
CC   well as tumour growth. The compositions of this invention exhibit
CC   cytostatic, virucidal, antibacterial and antiparasitic activities. This
CC   polynucleotide is a rat cDNA sequence that is necessary for viral
CC   infection, given in an exemplification of the invention.
XX
SQ   Sequence 982 BP; 324 A; 203 C; 201 G; 205 T; 0 U; 49 Other;

Query Match             3.4%;  Score 38.8;  Pred. No. 0.88;  Length 982;
Best Local Similarity  48.2%;
Matches   94;  Conservative    0;  Mismatches  101;  Indels    0;  Gaps    0;

Qy      610  GACTTATCATGGAGCCATTAAATTGGATGACAAAGATCCACAATTCAAAGACAACGTCAT  669
             |||  |||||||    |  | |||||||       |      |   |||||
```

```
Db      565  GTCTACACATGGAGAATAATANCTAGTTGACAAAAGCCANTTTAAGCAAGACATCTGCAA  624
             ||||||||| ||| ||| ||||  |||||||      ||
Qy      670  ACTGCTGAACAAGCACATTGACGCATACAACATTCCACCACAAAACATTCCCACCACAAAGGA  729
             |||||||  || || ||||
Db      625  TCCTCANCANCATGACATTGACTTTTAAATCACAGCCANNGACAACACTCCTAAANTTC  684
             ||||| |
Qy      730  CAAAAGAAAAAGACTGATGAAGCTCAGCCTTTGCCGCAGAGACAAAAGAGAGCAGCCCAC  789
             ||||||
Db      685  TAAAGACAACGGAATANTTAAATTCTGCAACCCNCAAAAAAAAAAAATCCACGGCANAA  744
             |||||
Qy      790  TGTGACTCCTTCTCC  804
             ||
Db      745  ATTGGGTTTTTTTCC  759

RESULT 15
ADA71938/c
ID   ADA71938/c standard; DNA; 2000 BP.
XX
AC   ADA71938;
XX
DT   20-NOV-2003 (first entry)
XX
DE   Rice gene, SEQ ID 5263.
XX
KW   Plant; bacterial infection; fungal infection; viral infection; rice;
KW   gene; ds.
XX
OS   Oryza sativa.
XX
PN   WO2003000898-A1.
XX
PD   03-JAN-2003.
XX
PF   22-JUN-2001; 2001WO-IB001105.
XX
PR   22-JUN-2001; 2001WO-IB001105.
XX
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI   Chang H,  Chen W,  Cooper B,  Glazebrook J,  Goff SA,  Hou Y;
PI   Katagiri F,  Quan S,  Tao Y,  Whitham S,  Xie Z,  Zhu T,  Zou G;
XX
DR   WPI; 2003-175290/17.
XX
PT   Identifying at least one gene involved in plant resistance or response to
PT   pathogenic infection for conferring resistance or tolerance to a plant to
PT   bacterial, fungal or viral infection by determining or detecting plant
PT   gene expression.
XX
PS   Claim 27; SEQ ID NO 5263; 899pp; English.
XX
CC   The present invention relates to a method (M1) for identifying genes
CC   involved in plant resistance or response to pathogenic infection. M1
CC   comprises identifying a gene whose expression is significantly altered in
CC   the incompatible interaction of plant gene expression relative to
CC   expression of the gene in an uninfected plant, in a mutant plant that
CC   does not express a gene associated with response to pathogenic infection,
CC   or in a corresponding incompatible or compatible interaction. (M1) is
CC   useful for conferring resistance to resistance or tolerance to a plant to
CC   bacterial, fungal or viral infection. The present sequence was used to
CC   illustrate the invention.
XX
SQ   Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

Query Match             3.4%;  Score 38.6;  DB 8;  Length 2000;
Best Local Similarity  10.2%;  Pred. No. 1.4;
Matches   67;  Conservative  299;  Mismatches  278;  Indels   11;  Gaps    3;

Qy      314  AAAGTTTCTGGTAAAGGCCAACAACAACAAGGCCAAACTGTCACTAAGAAATCTGCTGCT  373
             :::::| :||  :|:  :|: ||::  :||:|:: ||:: ||||: ::::: ::::
Db      853  MWWRYTMYTYCYAMTCAKCCKYKMAMTKWWTTWACAWRATSWRWRAMAGMRWKRYKMKRAY  794
```

```
Qy  374  GAGGCATCTAAAAAGCCTCGCCAAAAACGTACTGCCACAAAACAGTACAACGTCACTCAA  433
              ::  :  |:|::  :                   ::::|:  |::::::|  :  ::
Db  793  WWRWWRCWKAGWARWMKSRYRWKWKKYATRYYWKMWAMTWWWSWRRWKSYRMWSGMGRMR  734

Qy  434  GCATTTGGGAGACGTGGTCCAGAACAAACCCAAGAAATTTCG--GGGACCAAGACCTAA  491
              :|::  ::::::  ::|::  ::|::  :|  ::::::  :|::::  :  ::::|:
Db  733  WSAWRYCSRMKCAKTKYASSARWTKRAKRSYRYRRRWTYWKRKGWTYRYRYWRSCRMTRAR  674

Qy  492  TCAGACAAGGAACTGATTACAAACATTGGCCGCAAATTGCACAATTGCTCCAAGTGCCT  551
              :::  :|  ::  |    |:::::::::  :  ::::  |::|  :  ::  :  ::::
Db  673  MSKRRKWAGASMKSCWMYWRGARSMWYSKYSCSAKCCKKTRIYMTSSYMSTGMYGMYSSYK  614

Qy  552  CTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCACACCTTCGGGAACATGGCTGA  611
              :  :  :|::|:::  |:::  |::|  |:|::::  :  :  :::  :|
Db  613  SMSWTSKMSYMGKMTCTMYTSMKGSTRRSKMGRWSGMSRMYMRWWKKQMRKRKYMRYMKWK  554

Qy  612  CTTATCATGGAGCCATTAAATTGGATGACAAAGATCCACAATTCAAAGACAACGTCATAC  671
              ||:|:  |:  :::|:  |::::|:  ::|  :::::::  :  |:::::  :  :|:::
Db  553  CTWRRCMCYRWGYTMYTTSRSRMMYTGRYKARYTSKRYMWYKYRKYCWYYYYGMYMKCS  494

Qy  672  TGCTGAACAAGCACATTGACGCATACAAAACATTCCCACCACAGAGCCTAAAAAGGACA  731
              :::::  :::::  :  :|:::  ::  :  ::  ::  :  :  ::  :::::::
Db  493  Y----MMRYGYCKACKKCCYAMCWKAAYSGMMMYWYRKYSKWMRMSTKYMWSMWYKKCR  439

Qy  732  AAAAGAAAAAGACTGATGAAGCTCAGCCTTGCGCAGAGACAAAAGAGCAGCCCCACTG  791
              :  |:|::  |::  :  ::  :::  :::::  :  ::  ::::::::
Db  438  SMKYG----AKGCYGCKMWTYCSYGYMKWYTYMGSYKYSRCYKYMRMYMYKGWMYMMYYS  383

Qy  792  TGACTCTTCTTCCTCGCGGCTGACATGGATGATTTCTCCAGACAACTTCAAAATTCCATGA  851
              ::|:::::  :  |:|::  ::::  :|::::  |::  :  :  :::::::
Db  382  AYSSMMTWYYYAKYWYWYKRGTMSWYGKSYKKKYCTWWCYMKCMRCYRWRKMMRKKT  323

Qy  852  GTGGAGCTTCTGCTTCAACTCAGGCATAAACACTCATGATGACCACCAAGGCAGAT  911
              ::::::  :  :::  ::  ::::  :::::  ::  :  |::::::::::::
Db  322  KYSKRCYCWRYATCYWCCCYRKRGWYSRRSMMRTAGKWKMRSWSRWCRSYSWYKMYKRMW  263

Qy  912  GGGCTATGTAAACGTTTTCGCAATTCCGTTACGATACATAGTCTACTCTTGTGC  966
              :::::::  |::  |::|  :  |:::::  :  ||:::  :  |  ||::
Db  262  KKSYYMSYGWARSSGTWSRSAAKRTYKGYSTSRRAKMMRACRMYSACRRYSRTSY  208
```

Search completed: May 15, 2005, 11:25:14
Job time : 719.805 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 11:11:51 ; Search time 249.39 Seconds
                                      (without alignments)
                                      7453.434 Million cell updates/sec

Title: US-10-764-075-1
Perfect score: 1136
Sequence: 1 aggcacgtacggtgcaa........tgaataatgctagggagagc 1136

Scoring table: IDENTITY_NUC
                Gapop 10_0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_NA:*
          1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
          2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
          3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
          4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
          5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
          6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

   Pred. No. is the number of results predicted by chance to have a
   score greater than or equal to the score of the result being printed,
   and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match Score	Query Length	DB	ID	Description
c 1	43	3.8	1141	4	US-09-806-708B-22	Sequence 22, Appl
c 2	39.4	3.5	2889	4	US-09-248-796A-5056	Sequence 5056, App
c 3	38.6	3.4	7218	1	US-08-232-463-14	Sequence 14, Appl
c 4	37.4	3.3	13489	4	US-09-949-016-15911	Sequence 15911, A
c 5	36	3.2	601	4	US-09-949-016-53337	Sequence 53337, A
c 6	36	3.2	110402	4	US-09-949-016-17295	Sequence 17295, A
c 7	36	3.2	110403	4	US-09-949-016-12741	Sequence 12741, A
c 8	36	3.2	139936	4	US-09-949-016-11782	Sequence 11782, A
c 9	36	3.2	139952	4	US-09-949-016-13280	Sequence 13280, A
10	35.8	3.2	1450	4	US-09-620-312D-1058	Sequence 1058, Ap
11	35.8	3.2	1862	3	US-09-336-643A-11	Sequence 11, Appl
12	35.6	3.1	601	4	US-09-949-016-53338	Sequence 53338, A
13	34.8	3.1	601	4	US-09-949-016-102602	Sequence 102602,
14	34.8	3.1	601	4	US-09-949-016-102714	Sequence 102714,
15	34.8	3.1	6755	3	US-08-931-999-4	Sequence 4, Appli
c 16	34.8	3.1	678533	4	US-09-949-016-14577	Sequence 14577, A
c 17	34.8	3.1	678533	4	US-09-949-016-14578	Sequence 14578, A
18	34.6	3.0	320	4	US-09-621-976-2771	Sequence 2771, Ap
19	34.6	3.0	326	4	US-09-513-999C-10231	Sequence 10231, A
20	34.6	3.0	345	4	US-09-513-999C-10233	Sequence 10233, A
21	34.6	3.0	423	4	US-09-621-976-2768	Sequence 2768, Ap
22	34.6	3.0	424	4	US-09-621-976-2767	Sequence 2767, Ap
23	34.6	3.0	436	4	US-09-621-976-2766	Sequence 2766, Ap
24	34.6	3.0	504	4	US-09-621-976-16332	Sequence 16332, A
25	34.6	3.0	556	4	US-09-621-976-2769	Sequence 2769, Ap
26	34.6	3.0	1278	3	US-09-134-001C-2817	Sequence 2817, App
27	34.6	3.0	2175	4	US-09-710-279-967	Sequence 967, App
c 28	34.6	3.0	4216	4	US-09-710-279-3544	Sequence 3544, Ap
29	34.4	3.0	1506	4	US-09-107-433-269	Sequence 269, App
30	34.4	3.0	1632	4	US-09-581-110-2008	Sequence 2008, Ap
31	34.4	3.0	2335	3	US-08-961-527-275	Sequence 275, App
32	34.2	3.0	601	4	US-09-949-016-91446	Sequence 91446, A
33	34.2	3.0	143155	4	US-09-949-016-11925	Sequence 11925, A
34	34.2	3.0	143164	4	US-09-949-016-14368	Sequence 14368, A
35	34.2	3.0	143173	4	US-09-949-016-14513	Sequence 14513, A
36	34	3.0	320	3	US-09-030-607-224	Sequence 224, App
37	34	3.0	320	3	US-09-439-313-224	Sequence 224, App
38	34	3.0	320	3	US-09-352-616A-224	Sequence 224, App
39	34	3.0	320	3	US-09-233-149A-224	Sequence 224, App
40	34	3.0	320	4	US-09-159-812-224	Sequence 224, App
41	34	3.0	320	4	US-09-635-215-224	Sequence 224, App
42	34	3.0	320	4	US-09-685-166A-224	Sequence 224, App
43	34	3.0	320	4	US-09-115-453-224	Sequence 224, App
44	34	3.0	320	4	US-09-688-489-224	Sequence 224, App
45	34	3.0	320	4	US-09-679-426-224	Sequence 224, App

ALIGNMENTS

RESULT 1
US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAE1 promoters
US-09-806-708B-22

Query Match   3.8%; Score 43; DB 4; Length 1141;
Best Local Similarity 11.7%; Pred. No. 0.0059;
Matches   36; Conservative 142; Mismatches 130; Indels   1; Gaps   1;

Qy   739 AAAGACTGATGAAGCTCAGCCTTTGCCGCAGAGACAAAAGAAGCAGCGCCACTGTGACTCT 798
          ::::::.:::: .:|| ::::: .:::::. .::: ::.: :::||.::::::
Db   451 MWMRRAWKMWNMAWCRRAYCCNNNNNRACVWHCHKMRWTWKYMWKAACNNNNBKAMYMR 392

Qy   799 TCTTCCTGCGGCTGACATGGATGATTTCTCCAGACAACTTCAAAATTCCATGAGTGGAGC 858
          :::: .:::::: :: .:::: :::::::: .:|::::::||::: : :::::::::::
Db   391 VAWMMYSRDTTNTDWMMTSDWBWHWYTVDYTMRAWNNNNNNNRBCKTTSWMWWMDHM 332

Qy   859 TTCTGCTGATTCAACTCAGGCATAAACTCATGATGACCATAAAACTCAAGGCAGATGGGCTAT 918
          ||:.:: :|:|| ::|:: ::::: :::.|
Db   331 NTHCTYGNNTWGSAYBMAAMSMWAAGASNBVTYNWCWRMTYMGKTMTNNNNNNKAWYYRT 272

Qy   919 GTAAACGTTTTCGCAATTCCGTTTACGAT-ACATAGTCTACTCCTTGTCGCAGAATGAATTC 977
          |:.|:.::::.: :|:::::: .::.: .:.::::.:||
Db   271 KTVAWCNNRYYYDTAVWTBKRNYKYCYAYBWYYBMYMGKHHWBWWRRABHRSWNMWWVKC 212

Qy   978 TCGTAACTAAACAGCACAAGTAGGTTTAGTTAACTTTAATCTCACATAGCAATCTTTAAT 1037
          :::: ::.:::::::::: ::||: :||::::: :: .: ::: ::: .:|| :::: :::
Db   211 RNKYMVSWHYHAMRYBKWABAVGCNNNWKDRMAHHHWCATNNNMMWWYAYMHMHKKGK 152

Qy   1038 CAATGTGTA 1046
          ||:.:||

```
Db    151 AAWTNNKTA 143
```

RESULT 2
US-09-248-796A-5056
; Sequence 5056, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 5056
; LENGTH: 2889
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-5056

```
Query Match             3.5%;  Score 39.4; DB 4; Length 2889;
Best Local Similarity  56.6%;  Pred. No. 0.15;
Matches    73;  Conservative  73;  Mismatches  56;  Indels  0;  Gaps  0;

Qy   619 TGGAGCCATTAAATTGGATGACAAAGATCCACAATTCAAAGACAACGTCATACTGCTGAA 678
         || |||  |||||| || ||||| ||  |||||   | ||| |    |
Db  1674 TCGAGATATTACAGTGGATGAAAAACAATCAACCCTCCAAAAAACACAAAAGAGAACCGAC 1733

Qy   679 CAAGCACATTGACGCCATACACAAAACATTCCCACCACCAAAACATTCCCACCACCAAAAGAA 738
         ||||||||||||||   |||||  ||||  |||||||
Db  1734 TAAGGACATTGACACTGAAATCGAGAGAAACATAAATGCACTGTCAAAGTTAAAAATAT 1793

Qy   739 AAAGACTGA 747
         |||
Db  1794 CGCAGCTGA 1802
```

RESULT 3
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991

; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pT2gpt-F1s
US-08-232-463-14

```
Query Match             3.4%;  Score 38.6; DB 1; Length 7218;
Best Local Similarity   3.9%;  Pred. No. 0.46;
Matches   154;  Conservative  195;  Mismatches  154;  Indels  0;  Gaps  0;

Qy   422 AACGTCACTCAAGCATTTGGGAGACGTGGTCCAGAACAAACCCAAGGAATTTCGGGGAC 481
         |||| ||| ||  |||||||  |||||
Db  1457 AAAGAGATAGAAGAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1398

Qy   482 CAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAATTGCACAATTTGCT 541
         ::::  ::  :: :||   | :::: ::::  :: ::::
Db  1397 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1338

Qy   542 CCAAGTGCCTCGCATTCTTTGGAATGTCACGCATTGGCATGGAGTCACACCTTCGGGA 601
         ::   : ::::   :::::
Db  1337 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1278

Qy   602 ACATGGCTGACTTATCATGGAGCCATTAAATTGGATGACAAAGATCCACAATTCAAAGAC 661
         : ::  :  :::::: :: ::  :::: ::::
Db  1277 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1218

Qy   662 AACGTCATACTGCTGACAACAAGCACATTGACGCGCATACAAAACATTCCCACCACCAGAGCCT 721
         ::  ::  :  :::: ::  ::::: :::::
Db  1217 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1158

Qy   722 AAAAGGACAAAAGAAAAGACTGATGAAGCTCAGCCTTTGCCGCGAGAGACAAAAGAAG 781
         ::::::::  ::::::::::  :: :::::::::
Db  1157 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1098

Qy   782 CAG 784
         ::
Db  1097 RRR 1095
```

RESULT 4
US-09-949-016-15911/c
; Sequence 15911, Application US/0994916
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15911
; LENGTH: 13489
; TYPE: DNA

```
;  ORGANISM: Human
US-09-949-016-15911

Query Match             3.3%;       Score 37.4; DB 4; Length 13489;
Best Local Similarity  52.2%;       Pred. No. 1.6;
Matches   83; Conservative    0; Mismatches   76; Indels    0; Gaps    0;

Qy    583 GGAAGTCACACCTTCGGGAACATGGCTGACTTATCATGGAGCCATTAAATTGGATGACAA 642
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11661 GGAAGTAACACGTCTGGCACCAGAGTTGACTTTAATTGTCTCAAAGCGCCGCTGAAGCAA 11602

Qy    643 AGATCCACAATTCAAAGACAGACACGTCATACTGCTGACAAGCACATTGACGCATACAAAC 702
          ||||||||||| ||||| |||||||||||| |||| |  ||  |||  |||||| | ||||
Db  11601 AAACCATGATGATAAAACATTCTGCTTTCTTTTCTTTTTTACAACCCACGAACGGCAAAAAAAA 11542

Qy    703 ATTCCCACAACAGAGCCTAAAAAGGACAAAAAGAAAAA 741
          ||| |||| |||||||||| |||| || ||
Db  11541 AAAAAAAAAACAAAACAAAAAGAAAACAACAACAA 11503


RESULT 5
US-09-949-016-53337
;  Sequence 53337, Application US/09949016
;  Patent No. 6812339
;  GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE: 2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 53337
;  LENGTH: 601
;  TYPE: DNA
;  ORGANISM: Human
US-09-949-016-53337

Query Match             3.2%;       Score 36; DB 4; Length 601;
Best Local Similarity  54.5%;       Pred. No. 0.7;
Matches   72; Conservative    0; Mismatches   60; Indels    0; Gaps    0;

Qy    326 AAAGGCCAACAACAACAAGGCCAAACTGTCACTAAGAAATCTGCTGCTGAGGCATCTAAA 385
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db      4 AAGAGTCAATACCCACTGAAGCCACCATTCACGCGAGGAGCAACAAAAGAACCCTCCAAC 63

Qy    386 AAGGCCTCGCCAAAAACGTACTGCCACAAAACAGTACAACGTCACTCAAGCCATTTGGGAGA 445
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     64 AAGCAAGGGCACTGAAAGTACTGAATCAAACTTAAAGCACTCACACGAGCAAAGGGGAAAA 123

Qy    446 CGTGGTCCAGAA 457
          ||||| || ||
Db    124 CATGGTACAAAA 135


RESULT 6
US-09-949-016-17295/c
;  Sequence 17295, Application US/09949016
;  Patent No. 6812339
;  GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE: 2000-04-14
```

```
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 17295
;  LENGTH: 110402
;  TYPE: DNA
;  ORGANISM: Human
US-09-949-016-17295

Query Match             3.2%;       Score 36; DB 4; Length 110402;
Best Local Similarity  50.4%;       Pred. No. 16;
Matches  114; Conservative    0; Mismatches  110; Indels    2; Gaps    1;

Qy    652 ATTCAAAGACAACGTCATACTGCTGAACAAGCACATTGACGCATACAAAACATTCCACC 711
          ||| ||||| ||||     |  ||||  ||  ||  ||  ||  ||  |||| ||
Db  16399 ATATATATATACACATATACACACACATCTATCCTATTATTCTCTCCCTCT 16340

Qy    712 AACAGAGCCTAAAAAGGACAAAAAGAAAAAGACTGAAGCTTCAGCCTTGCCGCGAGAG 771
          ||||||   ||||| || ||||| ||||| |||  ||  ||  |||| |  | |||||
Db  16339 AGAGAACCCTAATACAATCATACACAAAATTTCTTAACCCAAAAGTGTTAGTGAAGAA 16280

Qy    772 ACAAAAGAAGCAGCCCACCTGTGACTCTTCTTCCTCGCGGTCGACGATGAGTTCTCCAG 831
          || ||||| ||  ||    ||  |  ||||   ||  | | || |||  | |||||||
Db  16279 ATACGAGGAGAATGTTAATGCTACTCT--TTCCTGAGCTGACCTGATTCAACTCAG 16222

Qy    832 ACAACTTCAAAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAG 877
          ||  || |||| ||     | |     ||  |||   |||
Db  16221 AGGTAGTTATTTTTCAGGAGTGGATCGGATCCAGCTGCAGACCTAAGCTAG 16176


RESULT 7
US-09-949-016-12741/c
;  Sequence 12741, Application US/09949016
;  Patent No. 6812339
;  GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE: 2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 12741
;  LENGTH: 110403
;  TYPE: DNA
;  ORGANISM: Human
US-09-949-016-12741

Query Match             3.2%;       Score 36; DB 4; Length 110403;
Best Local Similarity  50.4%;       Pred. No. 16;
Matches  114; Conservative    0; Mismatches  110; Indels    2; Gaps    1;

Qy    652 ATTCAAAGACAACGTCATACTGCTGAACAAGCACATTGACGCATACAAAACATTCCACC 711
          ||| ||||| ||||     |  ||||  ||  ||  ||  ||  ||  |||| ||
Db  16406 ATATATATATACACATATACACACACATCTATCCTATTATTCTCTCCCTCT 16347

Qy    712 AACAGAGCCTAAAAAGGACAAAAAGAAAAAGACTGAAGCTTCAGCCTTGCCGCGAGAG 771
          ||||||   ||||| || ||||| ||||| |||  ||  ||  |||| |  | |||||
Db  16346 AGAGAACCCTAATACAATCATACACAAAATTTCTTAACCCAAAAGTGTTAGTGAAGAA 16287

Qy    772 ACAAAAGAAGCAGCCCACCTGTGACTCTTCTTCCTCGCGGTCGACGATGAGTTCTCCAG 831
```

```
Db  16286 ATACGAGAGGAATGTTAATGCTACTCT--TTCCTGAGGCTGACAAAGACAAAGACAAGTA 16229
              ||||  ||||  |||||||||  ||||||  |||||| |||||| ||||||||||| |
Qy    832 ACAACTTCAAAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAG 877
                     |||||||||||||  |||| ||||| ||  || ||
Db  16228 AGGTAGTTATTTTTTCAGGAGTGGATCAGCTGGAGCTGAGACCTAAGCTAG 16183

RESULT 8
US-09-949-016-11782/c
; Sequence 11782, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;   APPLICANT: VENTER, J. Craig et al.
;   TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;   TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;   FILE REFERENCE: CL001307
;   CURRENT APPLICATION NUMBER: US/09/949,016
;   CURRENT FILING DATE: 2000-04-14
;   PRIOR APPLICATION NUMBER: 60/241,755
;   PRIOR FILING DATE: 2000-10-20
;   PRIOR APPLICATION NUMBER: 60/237,768
;   PRIOR FILING DATE: 2000-10-03
;   PRIOR APPLICATION NUMBER: 60/231,498
;   PRIOR FILING DATE: 2000-09-08
;   NUMBER OF SEQ ID NOS: 207012
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11782
;   LENGTH: 139936
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(139936)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11782

Query Match             3.2%;  Score 36;  DB 4;  Length 139936;
Best Local Similarity   54.5%;  Pred. No. 18;
Matches  72;  Conservative  0;  Mismatches  60;  Indels  0;  Gaps  0;

Qy    326 AAAGGCCAACAACAACAAGGCCAAACTGTCACTAAGAAATCTGCTGCTGAGGCATCTAAA 385
              |||||||||||||||||||||||||||  |||||||||| |||||||||| |||||||
Db  73881 AAGAGTCAATACCCACTGAAGCCACCATTCACGCAGGAGCAACAAAAGAACCCTCCAAC 73822

Qy    386 AAGGCCTCGCCAAAAACGTACTGCCACAAAAACAGTACAACGTCACTCAAGCATTTGGGAGA 445
              |||| ||||||||||||| |||||| |||||||||||||
Db  73821 AAGCAAGGCACTGAAAGTACTGAATCAAAACTTAAAGCACTCACACGAGCAAAGGGAAAA 73762

Qy    446 CGTGGTCCAGAA 457
              ||||||||||
Db  73761 CATGGTACAAAA 73750

RESULT 9
US-09-949-016-13280/c
; Sequence 13280, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;   APPLICANT: VENTER, J. Craig et al.
;   TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;   TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;   FILE REFERENCE: CL001307
;   CURRENT APPLICATION NUMBER: US/09/949,016
;   CURRENT FILING DATE: 2000-04-14
;   PRIOR APPLICATION NUMBER: 60/241,755
;   PRIOR FILING DATE: 2000-10-20
;   PRIOR APPLICATION NUMBER: 60/237,768
;   PRIOR FILING DATE: 2000-10-03
;   PRIOR APPLICATION NUMBER: 60/231,498
;   PRIOR FILING DATE: 2000-09-08
;   NUMBER OF SEQ ID NOS: 207012
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13280
;   LENGTH: 139952
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(139952)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13280

Query Match             3.2%;  Score 36;  DB 4;  Length 139952;
Best Local Similarity   54.5%;  Pred. No. 18;
Matches  72;  Conservative  0;  Mismatches  60;  Indels  0;  Gaps  0;

Qy    326 AAAGGCCAACAACAACAAGGCCAAACTGTCACTAAGAAATCTGCTGCTGAGGCATCTAAA 385
              |||||||||||||||||||||||||||  |||||||||| |||||||||| |||||||
Db  73896 AAGAGTCAATACCCACTGAAGCCACCATTCACGCAGGAGCAACAAAAGAACCCTCCAAC 73837

Qy    386 AAGGCCTCGCCAAAAACGTACTGCCACAAAAACAGTACAACGTCACTCAAGCATTTGGGAGA 445
              |||| ||||||||||||| |||||| |||||||||||||
Db  73836 AAGCAAGGCACTGAAAGTACTGAATCAAAACTTAAAGCACTCACACGAGCAAAGGGAAAA 73777

Qy    446 CGTGGTCCAGAA 457
              ||||||||||
Db  73776 CATGGTACAAAA 73765

RESULT 10
US-09-620-312D-1058
; Sequence 1058, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
;   APPLICANT: Tang, Y. Tom
;   APPLICANT: Liu, Chenghua
;   APPLICANT: Asundi, Vinod
;   APPLICANT: Zhang, Jie
;   APPLICANT: Ren, Feiyan
;   APPLICANT: Chen, Rui-hong
;   APPLICANT: Zhao, Qing A.
;   APPLICANT: Wehrman, Tom
;   APPLICANT: Xue, Aidong J.
;   APPLICANT: Yang, Yonghong
;   APPLICANT: Wang, Jian-Rui
;   APPLICANT: Zhou, Ping
;   APPLICANT: Ma, Yunqing
;   APPLICANT: Wang, Dunrui
;   APPLICANT: Wang, Zhiwei
;   APPLICANT: John Tillinghast
;   APPLICANT: Drmanac, Radoje T.
;   TITLE OF INVENTION: No. 6569662el Nucleic Acids and
;   TITLE OF INVENTION: Polypeptides
;   FILE REFERENCE: 784CIP2B
;   CURRENT APPLICATION NUMBER: US/09/620,312D
;   CURRENT FILING DATE: 2000-07-19
;   PRIOR APPLICATION NUMBER: 09/552,317
;   PRIOR FILING DATE: 2000-04-25
;   PRIOR APPLICATION NUMBER: 09/488,725
;   PRIOR FILING DATE: 2000-01-21
;   NUMBER OF SEQ ID NOS: 1105
;   SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 1058
;   LENGTH: 1450
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (397)..(1176)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1450)
;   OTHER INFORMATION: n = a,t,c or g
US-09-620-312D-1058
```

Query Match   3.2%;   Score 35.8; DB 4; Length 1450;
Best Local Similarity 69.0%;   Pred. No. 1.4;
Matches   49; Conservative   0; Mismatches 22; Indels   0; Gaps   0;

Qy   196 TTCAAGAAATTCAACTCCTGCGCAGCAGTAGGGGAAATTCTCCTGCTCGAATGGCTAGCGG 255
           ||| ||||||||||||||||| ||||| ||||| || | ||||||| | ||| |||||||
Db   707 TTCGAGAAAATCAACTTCTTGCACAAGAAGCAGAATTCTTTCAGCTCAAGGGACTGGCAG 766

Qy   256 AGGTGGTGAAA 266
           ||| |||||||
Db   767 AGGAAGTGAAA 777

RESULT 11
US-09-336-643A-11
; Sequence 11, Application US/09336643A
; Patent No. 6399761
; GENERAL INFORMATION:
; APPLICANT: Miller, Andrew P.
; APPLICANT: Curran, Mark Edward
; APPLICANT: Hu, Ping
; APPLICANT: Rutter, Marc
; APPLICANT: Wang, Jian-Wang
; TITLE OF INVENTION: No. 639976lel Human Potassium Channels
; FILE REFERENCE: SEQ-15P
; CURRENT APPLICATION NUMBER: US/09/336,643A
; CURRENT FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 60/076,687
; PRIOR FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: 60/116,448
; PRIOR FILING DATE: 1999-01-19
; PRIOR APPLICATION NUMBER: PCT/US99/03826
; PRIOR FILING DATE: 1999-02-22
; NUMBER OF SEQ ID NOS: 87
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (383)...(1157)
; OTHER INFORMATION: K+Hnov15
US-09-336-643A-11
; SEQ ID NO 11
; LENGTH: 1862
; TYPE: DNA
; ORGANISM: H. sapiens
; SOFTWARE: FastSEQ for Windows Version 4.0

Query Match   3.2%;   Score 35.8; DB 3; Length 1862;
Best Local Similarity 69.0%;   Pred. No. 1.6;
Matches   49; Conservative   0; Mismatches 22; Indels   0; Gaps   0;

Qy   196 TTCAAGAAATTCAACTCCTGCGCAGCAGTAGGGGAAATTCTCCTGCTCGAATGGCTAGCGG 255
           ||| ||||||||||||||||| ||||| ||||| || | ||||||| | ||| |||||||
Db   693 TTCGAGAAAATCAACTTCTTGCACAAGAAGCAGAATTCTTTCAGCTCAAGGGACTGGCAG 752

Qy   256 AGGTGGTGAAA 266
           ||| |||||||
Db   753 AGGAAGTGAAA 763

RESULT 12
US-09-949-016-53338
; Sequence 53338, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 53338
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-53338

Query Match   3.1%;   Score 35.6; DB 4; Length 601;
Best Local Similarity 53.8%;   Pred. No. 0.95;
Matches   71; Conservative   1; Mismatches 60; Indels   0; Gaps   0;

Qy   326 AAAGGCCAACAACAACAAGGCCAAACTGCTCACTAAGAAATCTGCTGCTGAGGCATCTAAA 385
           || ||| ||| | ||||| ||| || || ||||| || |||||||||||| ||||||
Db   192 AAGAGTCAATACCCACTGAAGCCACCATTCACGCAGGAGACAACAAAAGAACCCTCCAAC 251

Qy   386 AAGCCTCGCCAAAAACGTACTGCCACAAAAACAGTACAACGTCACTCAAGCATTGGGAGA 445
           |||||| | |||| ||||| | || || |||||||||||| ||| |||||| |||||
Db   252 AAGCAAGGCACTGAAGTACTGAATCAACAAACTTAAAGCACTCACCACGAGSAAAGGGAAA 311

Qy   446 CGTGGTCCAGAA 457
           |||| ||||||
Db   312 CATGGTACAAAA 323

RESULT 13
US-09-949-016-102602
; Sequence 102602, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 102602
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-102602

Query Match   3.1%;   Score 34.8; DB 4; Length 601;
Best Local Similarity 53.7%;   Pred. No. 1.7;
Matches   72; Conservative   0; Mismatches 62; Indels   0; Gaps   0;

Qy   653 TTCAAAGACAACGTCATCATCTGCTGACAAGCACATTGACGCATACAAAACATTCCACCA 712
           ||| ||| | | || | | ||| ||| ||||| | || | | ||| | | |||||||
Db   153 TTAAGAGCAGAAATGATAAAGCTGAAGAAACAATTAGTAAGCTTAAAAACAGGCTATTA 212

Qy   713 ACAGAGCCTAAAAAGGACAAAAGAAGAAAAAGACTGATGAAGCTCAGCCTTTGCCGCAGA 772
           | ||||| | || ||| ||| ||||||| | ||||| | || || | || |||||||
Db   213 AAAATACACAAGGACGACAAAAGAAGAAAAAGATAAAATGAATGACGCATGCCTACTAGA 272

Qy   773 CAAAGAAGCAGCC 786
           ||||| |||||
Db   273 TATAGAAAACAGCC 286

RESULT 14
US-09-949-016-102714
; Sequence 102714, Application US/09949016
; Patent No. 6812339

```
;  GENERAL INFORMATION:
;    APPLICANT: VENTER, J. Craig et al.
;    TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;    TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;    FILE REFERENCE: CL001307
;    CURRENT APPLICATION NUMBER: US/09/949,016
;    CURRENT FILING DATE: 2000-04-14
;    PRIOR APPLICATION NUMBER: 60/241,755
;    PRIOR FILING DATE: 2000-10-20
;    PRIOR APPLICATION NUMBER: 60/237,768
;    PRIOR FILING DATE: 2000-10-03
;    PRIOR APPLICATION NUMBER: 60/231,498
;    PRIOR FILING DATE: 2000-09-08
;    NUMBER OF SEQ ID NOS: 207012
;  SEQ ID NO 102714
;    SOFTWARE: FastSEQ for Windows Version 4.0
;    LENGTH: 601
;    TYPE: DNA
;    ORGANISM: Human
US-09-949-016-102714

Query Match             3.1%;    Score 34.8;  DB 4;  Length 601;
Best Local Similarity  53.7%;    Pred. No. 1.7;
Matches   72;  Conservative    0;  Mismatches  62;  Indels   0;  Gaps    0;

Qy      653 TTCAAAGAACAACGTCATACTGCTGAACAAGCACATTGACGCATACAAAACATTCCCACCA 712
                |||||||| |||  |||||| |||||| |||||||  || ||||||| |||||||||||
Db      153 TTAAGAGCAGAAATGATAAAGCTGAAGAACAATTAGTAAGCTTAAAAACAGGCTATTTA 212

Qy      713 ACAGAGCCTAAAAAGGACAAAAAGAACTGATGAAGCTCAGCCTTTGCCGCAGAGA 772
                ||||||||||||| ||||||| || |  ||||| |   || ||||||||||||
Db      213 AAAATACACAAAGGAGACAAAAGAAAAGAATTAAAATTAGAATTGACGCATGCCTACTAGA 272

Qy      773 CAAAAGAGCAGCC 786
                |  |||||||||
Db      273 TATAGAAAACAGCC 286


RESULT 15
US-08-931-999-4
;  Sequence 4, Application US/08931999
;  Patent No. 6043219
;  GENERAL INFORMATION:
;    APPLICANT: Iandolo, John J.
;    APPLICANT: Crupper, Scott S.
;    TITLE OF INVENTION: Broad Spectrum Chemotherapeutic Peptide
;    NUMBER OF SEQUENCES: 4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Hovey, Williams, Timmons & Collins
;      STREET: 2405 Grand Boulevard, Suite 400
;      CITY: Kansas City
;      STATE: Missouri
;      COUNTRY: U.S.A.
;      ZIP: 64108
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/931,999
;      FILING DATE:
;      CLASSIFICATION: 514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: US 08/710,561
;      FILING DATE: 19-SEP-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Collins, John M.
;      REGISTRATION NUMBER: 26,262
;      REFERENCE/DOCKET NUMBER: 25043-A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: 816/474-9050
;      TELEFAX: 816/474-9057
```

```
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 6755 base pairs
;      TYPE: nucleic acid
;      STRANDEDNESS: double
;      TOPOLOGY: unknown
;      MOLECULE TYPE: DNA (genomic)
;    HYPOTHETICAL: NO
;    ANTI-SENSE: NO
;    ORIGINAL SOURCE:
;      ORGANISM: Staphylococcus aureus
;      STRAIN: UT0007
US-08-931-999-4

Query Match             3.1%;    Score 34.8;  DB 3;  Length 6755;
Best Local Similarity  55.9%;    Pred. No. 7.2;
Matches   66;  Conservative    0;  Mismatches  52;  Indels   0;  Gaps    0;

Qy      677 AACAAGCACATTGACGCATACAAAACATTCCCACCAACAGAGCCTAAAAGGACAAAAAG 736
                ||||||||||  |||||| ||||||||||||| ||||| ||||  ||||||||||
Db     6508 AACAGCACAACAAAGCGAACAACACAACCCAAAAAAAAGAGAACAGAAGAAAACAAACACG 6567

Qy      737 AAAAAGACTGATGAAGCTCAGCCTTTGCCGCAGAGACAAAAGAGCAGCCCACTGTGA 794
                ||| | |  || || |||  ||||||||||| ||||||| ||| ||||   ||
Db     6568 AGAACAACAAAGAAGAAGGAGCAAAAGAACAACACAAAAAACACCAAGAACACGGGGA 6625
```

Search completed: May 15, 2005, 14:22:59
Job time : 253.39 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2005, 12:53:36 ; Search time 838.5 Seconds
            (without alignments)
            8291.679 Million cell updates/sec

Title: US-10-764-075-1
Perfect score: 1136
Sequence: 1 aggcatcgtatgggtgcaa.........tgaataatgctagggagagc 1136

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5562332 seqs, 3060109652 residues

Total number of hits satisfying chosen parameters: 11324664

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :   Published_Applications NA:*
    1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
    2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
    3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
    4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
    5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
    6: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
    7: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
    8: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
    9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
   10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
   11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
   12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
   13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
   14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
   15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
   16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
   17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
   18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
   19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
   20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
   21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
   22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1136	100.0	1136	18	US-10-764-075-1	Sequence 1, Appli
2	1136	100.0	1620	19	US-10-699-936-16	Sequence 16, Appli
3	1136	100.0	24774	19	US-10-889-447-3	Sequence 3, Appli
4	1136	100.0	28920	19	US-10-889-447-5	Sequence 5, Appli
5	1136	100.0	28920	19	US-10-889-447-6	Sequence 6, Appli
6	1136	100.0	29727	18	US-10-839-729-15	Sequence 15, Appli
7	1136	100.0	29727	19	US-10-827-757-1	Sequence 1, Appli
8	1136	100.0	29727	19	US-10-889-447-8	Sequence 8, Appli
9	1136	100.0	29727	19	US-10-699-936-1	Sequence 1, Appli
10	1136	100.0	29736	19	US-10-839-729-17	Sequence 17, Appli
11	1136	100.0	29736	19	US-10-889-447-9	Sequence 9, Appli

12	1136	100.0	29736	19	US-10-699-936-3	Sequence 3, Appli
13	1136	100.0	29742	18	US-10-839-729-16	Sequence 16, Appl
14	1136	100.0	29742	19	US-10-808-187-15	Sequence 15, Appl
15	1136	100.0	29742	19	US-10-808-187-16	Sequence 16, Appl
16	1136	100.0	29742	19	US-10-808-187-240	Sequence 240, App
17	1136	100.0	29742	19	US-10-808-187-737	Sequence 737, App
18 c	1136	100.0	29742	19	US-10-808-187-1108	Sequence 1108, Ap
19 c	1136	100.0	29742	19	US-10-808-187-1590	Sequence 1590, Ap
20 c	1136	100.0	29742	19	US-10-808-187-1965	Sequence 1965, Ap
21	1136	100.0	29742	19	US-10-889-447-10	Sequence 10, Appl
22	1136	100.0	29751	18	US-10-839-729-14	Sequence 14, Appl
23	1136	100.0	29751	19	US-10-856-529-1	Sequence 1, Appli
24	1136	100.0	29751	19	US-10-626-879-67	Sequence 67, Appl
25	1136	100.0	29751	19	US-10-889-447-1	Sequence 1, Appli
26	1136	100.0	29751	19	US-10-889-447-2	Sequence 2, Appli
27	1136	100.0	29751	19	US-10-699-936-2	Sequence 2, Appli
28	1134.4	99.9	29430	19	US-10-889-447-7	Sequence 7, Appli
29	1131.2	99.6	2304	19	US-10-699-936-2	Sequence 2, Appli
30	1131.2	99.6	2304	19	US-10-699-936-11	Sequence 11, Appli
31	1124	98.1	29291	19	US-10-889-447-4	Sequence 4, Appli
32 c	41.8	3.7	7359	15	US-10-311-455-1836	Sequence 1836, Ap
33 c	40.2	3.5	2543	18	US-10-363-345A-30203	Sequence 30203, A
34	40.2	3.5	2543	19	US-10-363-345-30204	Sequence 30204, A
35 c	40.2	3.5	2543	19	US-10-363-493A-30203	Sequence 30203, A
36	40.2	3.5	2543	19	US-10-363-493A-30204	Sequence 30204, A
37 c	39.2	3.5	9830	17	US-10-221-613-112	Sequence 112, App
38	39	3.4	481	18	US-10-357-930-57047	Sequence 57047, A
39	39	3.4	5284	18	US-10-723-860-7691	Sequence 7691, Ap
40 c	38.8	3.4	852	19	US-10-425-115-19938	Sequence 19938, A
41	38.2	3.4	2544	18	US-10-437-963-74109	Sequence 74109, A
42 c	38.2	3.4	6888	18	US-10-602-494-252	Sequence 252, App
43 c	38	3.3	8305	15	US-10-311-455-1541	Sequence 1541, Ap
44	37.8	3.3	2748	19	US-10-741-849-6128	Sequence 6128, Ap
45	37.6	3.3	438	18	US-10-437-963-89737	Sequence 89737, A

ALIGNMENTS

RESULT 1
US-10-764-075-1
; Sequence 1, Application US/10764075
; Publication No. US20040265796A1
; GENERAL INFORMATION:
;   APPLICANT: Briese, Thomas
;   APPLICANT: Lipkin, Ian W.
;   APPLICANT: Palacios, Gustavo
;   APPLICANT: Jabado, Omar
;   TITLE OF INVENTION: Methods and Kits for Detecting SARS-Associated Coronavirus
;   FILE REFERENCE: 5199-87
;   CURRENT APPLICATION NUMBER: US/10/764,075
;   CURRENT FILING DATE: 2004-01-23
;   PRIOR APPLICATION NUMBER: Provisional Application 60/463,704
;   PRIOR FILING DATE: 2003-04-17
;   NUMBER OF SEQ ID NOS: 16
;   SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
;   LENGTH: 1136
;   TYPE: DNA
;   ORGANISM: artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic nucleic acid sequence that includes the 3' non-coding
;   OTHER INFORMATION: region of the SARS-associated coronavirus genome and a portion o
;   OTHER INFORMATION: the N gene of the SARS-associated coronavirus genome
US-10-764-075-1

Query Match   100.0%; Score 1136; DB 18; Length 1136;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1136; Mismatches 0; Conservative 0; Indels 0; Gaps 0;

Qy   1 AGGCATCGTATGGGTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACATTGGCAC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 AGGCATCGTATGGGTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACATTGGCAC 60

RESULT 2
US-10-699-936-16
; Sequence 16, Application US/10699936
; Publication No. US20050095582A1
; GENERAL INFORMATION:
; APPLICANT: Gillim-Ross, Laura
; APPLICANT: Taylor, Jill
; APPLICANT: Scholl, David R.
; APPLICANT: Wentworth, David E.
; APPLICANT: Jollick, Joseph D.
; TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory
; TITLE OF INVENTION: Syndrome Coronavirus
; FILE REFERENCE: DHI-07986
; CURRENT APPLICATION NUMBER: US/10/699,936
; CURRENT FILING DATE: 2003-11-03
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16
; LENGTH: 1620
; TYPE: DNA
; ORGANISM: SARS-CoV ZJ-HZ01
US-10-699-936-16

Query Match   100.0%;   Score 1136;   DB 19;   Length 1620;
Best Local Similarity   100.0%;
Matches 1136;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACCATTGGCAC   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   461 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACCATTGGCAC 520

Qy   61 CCGGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGGAACAACATTGCC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   521 CCGGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGGAACAACATTGCC 580

Qy   121 AAAAGGCTTCTACGCGAGAGGGAAGCAGAGGCCGGCAGTCAAGCTCTTCTCGCTCCTCATC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   581 AAAAGGCTTCTACGCGAGAGGGAAGCAGAGGCCGGCAGTCAAGCTCTTCTCGCTCCTCATC 640

Qy   181 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGGCAGCAGTAGGGGAAATTCTCCTGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   641 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGGCAGCAGTAGGGGAAATTCTCCTGC 700

Qy   241 TCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTCGCCCTCGCGCTATTGCTCTAGACAGATTGAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   701 TCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTCGCCCTCGCGCTATTGCTCTAGACAGATTGAA 760

Qy   301 CCAGCTTGAGAGACAAAGTTTCTCGGTAAAGCCAAACTGTCACTAA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   761 CCAGCTTGAGAGACAAAGTTTCTCGGTAAAGCCAAACTGTCACTAA 820

Qy   361 GAAATCTGCTGCTCAAGCATCTAAAAAGCCCAAGGAAATTTCGGGGA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   821 GAAATCTGCTGCTCAAGCATCTAAAAAGCCCAAGGAAATTTCGGGGA 880

Qy   421 CAACGTCACTCACTCAAGCATTTGGGAGACGTGGTCCAGAACAAACAAACCCAAGGAAATTTCGGGGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   881 CAACGTCACTCACTCAAGCATTTGGGAGACGTGGTCCAGAACAAACAAACCCAAGGAAATTTCGGGGA 940

Qy   481 CCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAATTTGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   941 CCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAATTTGC 1000

Qy   541 TCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAGTCACACCTTCGGG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1001 TCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAGTCACACCTTCGGG 1060

Qy   601 AACATGGCTGACTTATCATCATGGAGCCATTAAATTGGATGACAAAGATCCACAATTCAAAGA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1061 AACATGGCTGACTTATCATCATGGAGCCATTAAATTGGATGACAAAGATCCACAATTCAAAGA 1120

Qy   661 CAACGTCATACTGCTGAACAAGCACACATTGACGCGATCAAGCCTATGTAAACGTTTTCGCAATT   720

First alignment column

```
Qy   181 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCTGGCGAGCAGTAGGGAAATTCCTCGC 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 23799 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCTGGCGAGCAGTAGGGAAATTCCTCGC 23858

Qy   241 TCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTCGCGGTATTGCTGCTAGACAGATTGAA 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 23859 TCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTCGCGGTATTGCTGCTAGACAGATTGAA 23918

Qy   301 CCAGCTTGAGAGCAAAGTTTCTGTAAAGGCCAACACACAAGGCCAACTGTCACTAA 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 23919 CCAGCTTGAGAGCAAAGTTTCTGTAAAGGCCAACACACAAGGCCAACTGTCACTAA 23978

Qy   361 GAAATCTGCTGCTGGAGGCATCTAAAAAGCCTGCCAAAAACGTACTGCCACAAAACAGTA 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 23979 GAAATCTGCTGCTGGAGGCATCTAAAAAGCCTGCCAAAAACGTACTGCCACAAAACAGTA 24038

Qy   421 CAACGTCACTCAAGCATTGGGAGACGTGGTCCAGAACAAAACCAAGGAAATTCGGGGA 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 24039 CAACGTCACTCAAGCATTGGGAGACGTGGTCCAGAACAAAACCAAGGAAATTCGGGGA 24098

Qy   481 CCAAGACCTAATCAGACAAGAGACTGATTACAAACATTGGCCGCAAATTGCACATTTGC 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 24099 CCAAGACCTAATCAGACAAGAGACTGATTACAAACATTGGCCGCAAATTGCACATTTGC 24158

Qy   541 TCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGTGAAGTCACACCTTCGGG 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 24159 TCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGTGAAGTCACACCTTCGGG 24218

Qy   601 AACATGGCTGACTTATCATGGAGCCATTAAATTGGATGACAAAGATCCACAATTCAAGA 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 24219 AACATGGCTGACTTATCATGGAGCCATTAAATTGGATGACAAAGATCCACAATTCAAGA 24278

Qy   661 CAACGTCATCATCTGCTGACAAGCACACATTGACGACATACAAAACATTCCCACCACAGAGCC 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 24279 CAACGTCATCATCTGCTGACAAGCACACATTGACGACATACAAAACATTCCCACCACAGAGCC 24338

Qy   721 TAAAAAGGACAAAAGAAAAAGAAAAAGACTGATGAAGCTCAGCCTTTGCCGCAGAGACAAAAGAA 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 24339 TAAAAAGGACAAAAGAAAAAGAAAAAGACTGATGAAGCTCAGCCTTTGCCGCAGAGACAAAAGAA 24398

Qy   781 GGAGCCCACTGTGACTCTTCTCTCCTGCGCGGCTGACATGGATGATTTCTCCAGACAACTTCA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 24399 GCAGCCCACTGTGACTCTTCTCTCCTGCGCGGCTGACATGGATGATTTCTCCAGACAACTTCA 24458

Qy   841 AAATTCCATGAGTGGGAGCTTCTGCTGATTCAACTCAGGCATAAACACTCATGATGACCAC 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 24459 AAATTCCATGAGTGGGAGCTTCTGCTGATTCAACTCAGGCATAAACACTCATGATGACCAC 24518

Qy   901 ACAAGGCAGATGGGCTATGTAAACGTTTCGCCAATTCCGTTTACGATACATAGTCTACTC 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 24519 ACAAGGCAGATGGGCTATGTAAACGTTTCGCCAATTCCGTTTACGATACATAGTCTACTC 24578

Qy   961 TTGTGCAGAATGAATTCTCGTAACTAAACAGACAAGTAGGTTTAGTTAACTTTAATCTC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 24579 TTGTGCAGAATGAATTCTCGTAACTAAACAGACAAGTAGGTTTAGTTAACTTTAATCTC 24638

Qy  1021 ACATAGCAATCTTAATCAATGTGTACAACTTCAAGGAGGACTTGAAAGAGCCACCACATTT 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 24639 ACATAGCAATCTTAATCAATGTGTAACATTGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 24698

Qy  1081 TCATCGAGGCCACGCGGAGTACGAGGGGTACAGTGAATAATGCTAGGGAGAGC 1136
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 24699 TCATCGAGGCCACGCGGAGTACGAGGGGTACAGTGAATAATGCTAGGGAGAGC 24754
```

Second alignment column

```
Db  1121 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         CAACGTCATCATCTGCTGACAACAAGCACACATTGACGCATACAAAAACATTCCCACCACAGAGCC 1180

Qy   721 TAAAAAGGACAAAAGAAAAAGAAAAAGACTGATGAAGCTCAGCCTTTGCCGCAGAGACAAAAGAA 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1181 TAAAAAGGACAAAAGAAAAAGAAAAAGACTGATGAAGCTCAGCCTTTGCCGCAGAGACAAAAGAA 1240

Qy   781 GCAGCCCACTGTGACTCTTCTCTCCTGCGCGGCTGACATGGATGATTTCTCCAGACAACTTCA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1241 GCAGCCCACTGTGACTCTTCTCTCCTGCGCGGCTGACATGGATGATTTCTCCAGACAACTTCA 1300

Qy   841 AAATTCCATGAGTGGGAGCTTCTGCTGATTCAACTCAGGCATAAACACTCATGATGACCAC 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1301 AAATTCCATGAGTGGGAGCTTCTGCTGATTCAACTCAGGCATAAACACTCATGATGACCAC 1360

Qy   901 ACAAGGCAGATGGGCTATGTAAACGTTTCGCCAATTCCGTTTACGATACATAGTCTACTC 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1361 ACAAGGCAGATGGGCTATGTAAACGTTTCGCCAATTCCGTTTACGATACATAGTCTACTC 1420

Qy   961 TTGTGCAGAATGAATTCTCGTAACTAAACAGCACAAGTAGGTTTAGTTAACTTTAATCTC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1421 TTGTGCAGAATGAATTCTCGTAACTAAACAGCACAAGTAGGTTTAGTTAACTTTAATCTC 1480

Qy  1021 ACATAGCAATCTTTAATCAATGTAACATTGTAACATTAGGGAGGACTTGAAAAGAGCCACCACATTT 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1481 ACATAGCAATCTTTAATCAATGTAACATTGTAACATTAGGGAGGACTTGAAAAGAGCCACCACATTT 1540

Qy  1081 TCATCGAGGCCACGCGGAGTACGAGGGGTACAGTGAATAATGCTAGGGAGAGC 1136
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1541 TCATCGAGGCCACGCGGAGTACGAGGGGTACAGTGAATAATGCTAGGGAGAGC 1596
```

Query Match 100.0%; Score 1136; DB 19; Length 24774;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     1 AGGCATCGTATGGGTTGCAACTGAACTGAGGGAGCCTTGAATACACACCCAAAGACCACCATTGGCAC 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 23619 AGGCATCGTATGGGTTGCAACTGAACTGAGGGAGCCTTGAATACACACCCAAAGACCACCATTGGCAC 23678

Qy    61 CCGCAATCTAATAACAATGTGCCACCGTGTACAACTTCCTCAAGGACAACAACATTGCC 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 23679 CCGCAATCCTAATAACAATGTGCCACCGTGCTACAACTTCCTCAAGGACAACAACATTGCC 23738

Qy   121 AAAAGGCTTCTCACGCGACCAGAGGAAGCAGAGGCGGCAGTCACCTTCTTCTCGGCTCCTCATC 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 23739 AAAAGGCTTCTCACGCGACCAGAGGAAGCAGAGGCGGCAGTCACCTTCTTCTCGGCTCCTCATC 23798
```

```
;  FILE REFERENCE: RTS-0685US
;  CURRENT APPLICATION NUMBER: US/10/889,447
;  CURRENT FILING DATE: 2004-07-12
;  PRIOR APPLICATION NUMBER: 60/486,670
;  PRIOR FILING DATE: 2003-07-12
;  NUMBER OF SEQ ID NOS: 241
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 5
;  LENGTH: 28920
;  TYPE: DNA
;  ORGANISM: SARS coronavirus isolate BJ03
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION:
;  OTHER INFORMATION: n is any nucleotide
US-10-889-447-5

Query Match          100.0%;  Score 1136;  DB 19;  Length 28920;
Best Local Similarity 100.0%;  Pred. No. 0;
Matches 1136;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACCATTGGCAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  27768 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACCATTGGCAC 27827

Qy     61 CCGCAATCCTAATAACAATGCTGCCACCGTCTACAACTTCCTCAAGGAACAACATTGCC 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  27828 CCGCAATCCTAATAACAATGCTGCCACCGTCTACAACTTCCTCAAGGAACAACATTGCC 27887

Qy    121 AAAAGGCTTCTACGCAGAGGGAAGCAGAGGCGGCAGTCAAGCCTCTTCTCGCTCCTCATC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  27888 AAAAGGCTTCTACGCAGAGGGAAGCAGAGGCGGCAGTCAAGCCTCTTCTCGCTCCTCATC 27947

Qy    181 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGGCAGCAGTAGGGGAAATTCTCCTGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  27948 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGGCAGCAGTAGGGGAAATTCTCCTGC 28007

Qy    241 TCGAATGGCTAGCCGAGGTGGTGAAACTGCCCTCGCGCTATTGCTAGACAGATTGAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28008 TCGAATGGCTAGCCGAGGTGGTGAAACTGCCCTCGCGCTATTGCTAGACAGATTGAA 28067

Qy    301 CCAGCTTGAGAGCAAAGTTTCTGGTAAAGGCCAACAACAACAAGGCCAAACTGTCACTAA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28068 CCAGCTTGAGAGCAAAGTTTCTGGTAAAGGCCAACAACAACAAGGCCAAACTGTCACTAA 28127

Qy    361 GAAATCTGCTCGCTGAGGCATCTAAAAAGCCTGCCACAAAACGTACTGCCACAAAACAGTA 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28128 GAAATCTGCTCGCTGAGGCATCTAAAAAGCCTGCCACAAAACGTACTGCCACAAAACAGTA 28187

Qy    421 CAACGTCACTCAAGCATTTGGGAGACGTGGTCCAGAAGAACGTCGTCCAGAAGAAATTTCGGGGA 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28188 CAACGTCACTCAAGCATTTGGGAGACGTGGTCCAGAAGAACCCAGAAGAAATTTCGGGGA 28247

Qy    481 CCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAAATTGCACAATTTGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28248 CCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAAATTGCACAATTTGC 28307

Qy    541 TCCAAGTGCCTCTGCATTCTTGGAATGTCACGCATTGGCATTGGAAGTCACACCTTCGGG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28308 TCCAAGTGCCTCTGCATTCTTGGAATGTCACGCATTGGCATTGGAAGTCACACCTTCGGG 28367

Qy    601 AACATGGCTGACTTATCATGGAGCCATTAAATTGGATGACAAAGATCCCACAATTCAAAGA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28368 AACATGGCTGACTTATCATGGAGCCATTAAATTGGATGACAAAGATCCCACAATTCAAAGA 28427

Qy    661 CAAGCGTCATACTGCTGAACAAGCACATTGACGCATACAAAACATTCCCACCAACAGAGCC 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28428 CAACGTCATACTGCTGAACAAGCACATTGACGCATACAAAACATTCCCACCAACAGAGCC 28487

Qy    721 TAAAAAGGACAAACAAAAAGAAAAAGACTGATGAAGCTCAGCCTTTGCCGCAGAGACAAAAGAA 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28488 TAAAAAGGACAAACAAAAAGAAAAAGACTGATGAAGCTCAGCCTTTGCCGCAGAGACAAAAGAA 28547
```

```
Qy    781 GCAGCCCACTGTGACTCTTCTTCCTCGCGGCTGACATGGATGATTCTCCAGACAACTTCA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28548 GCAGCCCACTGTGACTCTTCTTCCTCGCGGCTGACATGGATGATTCTCCAGACAACTTCA 28607

Qy    841 AAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAGGCATAAACACTCATGATGACCAC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28608 AAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAGGCATAAACACTCATGATGACCAC 28667

Qy    901 ACAAGGCAGATGGGCTATGTAAGTTTTCGCAATTCCGTTTACGATAACATAGTCTACTC 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28668 ACAAGGCAGATGGGCTATGTAAGTTTTCGCAATTCCGTTTACGATAACATAGTCTACTC 28727

Qy    961 TTGTGCAGAATGAATTCTCTAACTAAACAGCACAAGTAGTTTAGTTAACTTTAATCTC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28728 TTGTGCAGAATGAATTCTCTAACTAAACAGCACAAGTAGTTTAGTTAACTTTAATCTC 28787

Qy   1021 ACATAGCAATCTTTAATCAATGTGTAACATCAATGTGTAACATTAGGGACTTGAAAGAGCCACCACATTT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28788 ACATAGCAATCTTTAATCAATGTGTAACATTAGGGACTTGAAAGAGCCACCACATTT 28847

Qy   1081 TCATCGAGGCCACGCGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGGAGAGC 1136
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28848 TCATCGAGGCCACGCGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGGAGAGC 28903
```

```
RESULT 5
US-10-889-447-6
;  Sequence 6, Application US/10889447
;  Publication No. US2005007530TA1
;  GENERAL INFORMATION:
;  APPLICANT: Bennett, C. Frank
;  APPLICANT: Jain, Ravi
;  TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
;  FILE REFERENCE: RTS-0685US
;  CURRENT APPLICATION NUMBER: US/10/889,447
;  CURRENT FILING DATE: 2004-07-12
;  PRIOR APPLICATION NUMBER: 60/486,670
;  PRIOR FILING DATE: 2003-07-12
;  NUMBER OF SEQ ID NOS: 241
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 6
;  LENGTH: 28920
;  TYPE: DNA
;  ORGANISM: SARS coronavirus isolate BJ04
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION:
;  OTHER INFORMATION: n is any nucleotide
US-10-889-447-6

Query Match          100.0%;  Score 1136;  DB 19;  Length 28920;
Best Local Similarity 100.0%;  Pred. No. 0;
Matches 1136;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACCATTGGCAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  27768 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACCATTGGCAC 27827

Qy     61 CCGCAATCCTAATAACAATGCTGCCACCGTCTACAACTTCCTCAAGGAACAACATTGCC 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  27828 CCGCAATCCTAATAACAATGCTGCCACCGTCTACAACTTCCTCAAGGAACAACATTGCC 27887

Qy    121 AAAAGGCTTCTACGCAGAGGGAAGCAGAGGCGGCAGTCAAGCCTCTTCTCGCTCCTCATC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  27888 AAAAGGCTTCTACGCAGAGGGAAGCAGAGGCGGCAGTCAAGCCTCTTCTCGCTCCTCATC 27947

Qy    181 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGGCAGCAGTAGGGGAAATTCTCCTGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  27948 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGGCAGCAGTAGGGGAAATTCTCCTGC 28007

Qy    241 TCGAATGGCTAGCCGAGGTGGTGAAACTGCCCTCGCGCTATTGCTAGACAGATTGAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28008 TCGAATGGCTAGCCGAGGTGGTGAAACTGCCCTCGCGCTATTGCTAGACAGATTGAA 28067
```

```
; SEQ ID NO 15
; LENGTH: 29727
; TYPE: DNA
; ORGANISM: SARS Coronavirus
US-10-839-729-15

Query Match            100.0%;  Score 1136;  DB 18;  Length 29727;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 1136;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      1 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACACCATTGGCAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28506 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACACCATTGGCAC 28565

Qy     61 CCGCAATCCTAATAACAATGTCTGCCACCGTGCTACAACTTCCTCAAGGAACACATTGCC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28566 CCGCAATCCTAATAACAATGTCTGCCACCGTGCTACAACTTCCTCAAGGAACACATTGCC 28625

Qy    121 AAAAGGCTTCTACGCGAGGAAGCAGAGGGCGCAGTCAAGCCTCTTCTCGCTCCTCATC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28626 AAAAGGCTTCTACGCGAGGAAGCAGAGGGCGCAGTCAAGCCTCTTCTCGCTCCTCATC 28685

Qy    181 ACGTAGTCGCGGTAATTCAAGAAATTCAAGAAATTCAACTCGTCCCTCGCCGCTATTGAA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28686 ACGTAGTCGCGGTAATTCAAGAAATTCAAGAAATTCAACTCGTCCCTCGCCGCTATTGAA 28745

Qy    241 TCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTCGCCGCTATTGCTGTGCTAGACAGATTGAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28746 TCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTCGCCGCTATTGCTGTGCTAGACAGATTGAA 28805

Qy    301 CCAGCTTGAGAGCAAAGTTTCGGTAAAGGCCAACAACAAGGCCAAACTGCCACAAAACAGTA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28806 CCAGCTTGAGAGCAAAGTTTCGGTAAAGGCCAACAACAAGGCCAAACTGCCACAAAACAGTA 28865

Qy    361 GAAATCTGCTCTGAGGCATCTAAAAAGCCTGGTCCAGAGAACAAACCCAAGGAAATTTCGGGGA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28866 GAAATCTGCTCTGAGGCATCTAAAAAGCCTGGTCCAGAGAACAAACCCAAGGAAATTTCGGGGA 28925

Qy    421 CAACGTCACTCAAGCATTGGGAGCTCTGAGTCCAGAGAACTATTACAAAACAGTAAC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28926 CAACGTCACTCAAGCATTGGGAGCTCTGAGTCCAGAGAACTATTACAAAACAGTAAC 28985

Qy    481 CCAAGACCTAATCAGACAAGGAACTGATTACAAAACATTGGCCGCAAATTGCACAATTGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28986 CCAAGACCTAATCAGACAAGGAACTGATTACAAAACATTGGCCGCAAATTGCACAATTGC 29045

Qy    541 TCCAAGTGCCCTCGATTCTTCTTGGAATGTCGCATGGAGTCACACCTTCGGG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29046 TCCAAGTGCCCTCGATTCTTCTTGGAATGTCGCATGGAGTCACACCTTCGGG 29105

Qy    601 AACATGCTGACTTATCATCATGGAGCCATTAGGGAGGACTTGAAAGAGCCCACCACATTT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29106 AACATGCTGACTTATCATCATGGAGCCATTAGGGAGGACTTGAAAGAGCCCACCACATTT 29165

Qy    661 CAACGTCATCACTGCTGAACAGCACATTGACGACTCAGCACAAAACATTCCCAACAGAGCC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29166 CAACGTCATCACTGCTGAACAGCACATTGACGACTCAGCACAAAACATTCCCAACAGAGCC 29225

Qy    721 TAAAAGGACAAAAAGGCTTGCCGCAGCTCAGGAGAGAGTCAGCCTTTGCCGCAGACAACTTCA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29226 TAAAAGGACAAAAAGGCTTGCCGCAGCTCAGGAGAGAGTCAGCCTTTGCCGCAGACAACTTCA 29285

Qy    781 GCAGCCCACTGTGACTCTTCTTCCTGCGGGCTGACATGGATGATTTCTCCAGACAACTTCA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29286 GCAGCCCACTGTGACTCTTCTTCCTGCGGGCTGACATGGATGATTTCTCCAGACAACTTCA 29345

Qy    841 AAATTCCATGAGTGGAGCTTCGCTTCGATTCAACTCAACACTCATGATGACCAC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29346 AAATTCCATGAGTGGAGCTTCGCTTCGATTCAACTCAACACTCATGATGACCAC 29405

Qy    901 ACAAGGCAGATGGGCTATGTAAACGTTTTCGCAATTCGTAAACGTTTTCGCAATTCGTAAACGTCTACTC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29406 ACAAGGCAGATGGGCTATGTAAACGTTTTCGCAATTCGTAAACGTTTTCGCAATTCGTAAACGTCTACTC 29465

Qy    301 CCAGCTTGAGAGCAAAGTTTCGGTAAAGGCCAACAACAAGGCCAAACTGTCACTAA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28068 CCAGCTTGAGAGCAAAGTTTCGGTAAAGGCCAACAACAAGGCCAAACTGTCACTAA 28127

Qy    361 GAAATCTGCTCTGAGGCATCTAAAAAGCCTGCCACAAAACAGTACTGCCACAAAACAGTA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28128 GAAATCTGCTCTGAGGCATCTAAAAAGCCTGCCACAAAACAGTACTGCCACAAAACAGTA 28187

Qy    421 CAACGTCACTCAAGCATTGGGAGACGTGGTCCAGAACAACCCAAGGAAATTTCGGGGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28188 CAACGTCACTCAAGCATTGGGAGACGTGGTCCAGAACAACCCAAGGAAATTTCGGGGA 28247

Qy    481 CCAAGACCTAATCAGACAAGGAACTGATTACAAAACATTGGCCGCAAATTGCACAATTGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28248 CCAAGACCTAATCAGACAAGGAACTGATTACAAAACATTGGCCGCAAATTGCACAATTGC 28307

Qy    541 TCCAAGTGCCCTCGATTCTTCTTTGGAATGTCGCATGGAGTCACACCTTCGGG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28308 TCCAAGTGCCCTCGATTCTTCTTTGGAATGTCGCATGGAGTCACACCTTCGGG 28367

Qy    601 AACATGGCTGACTTATCATCATGGAGCCATTAAATTGGATGACAAAGATCCACAATTCAAAGA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28368 AACATGGCTGACTTATCATCATGGAGCCATTAAATTGGATGACAAAGATCCACAATTCAAAGA 28427

Qy    661 CAACGTCATACTGCTGAACAGCACATTGACGACGCATACAAAACATTCCCACCACAGAGCC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28428 CAACGTCATACTGCTGAACAGCACATTGACGACGCATACAAAACATTCCCACCACAGAGCC 28487

Qy    721 TAAAAGGACAAAAGAAAAAGAAAAGAAAAGAAGACAAAAAAGAAAAAGAAAGAA 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28488 TAAAAGGACAAAAGAAAAAGAAAAGAAAAGAAGACAAAAAAGAAAAAGAAAGAA 28547

Qy    781 GCAGCCCACTGTATGTGACTCTTCCTGCGCTGACTGATTCAACTCCGTTACCGTGATCGACCAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28548 GCAGCCCACTGTATGTGACTCTTCCTGCGCTGACTGATTCAACTCCGTTACCGTGATCGACCAC 28607

Qy    841 AAATTCCATGAGTGGAGCTTCGCTATGTAAACGTTTTCGCAATTGCACATAAAACACTCATGATGACCAC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28608 AAATTCCATGAGTGGAGCTTCGCTATGTAAACGTTTTCGCAATTGCACATAAAACACTCATGATGACCAC 28667

Qy    901 ACAAGGCAGATGGGCTATGTGACTCTTCCTGCGGGCTATGTGACTCTTCCTGCGCAATTCGTGTTTTAGTGTCTACTC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28668 ACAAGGCAGATGGGCTATGTGACTCTTCCTGCGGGCTATGTGACTCTTCCTGCGCAATTCGTGTTTTAGTGTCTACTC 28727

Qy    961 TTGTGCAGAATGAATTCTCGTAACTAAAACAGCACACAAGTAGGTTTAGTGTTTAATCTC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28728 TTGTGCAGAATGAATTCTCGTAACTAAAACAGCACACAAGTAGGTTTAGTGTTTAATCTC 28787

Qy   1021 ACATAGCAATCTTTAATCAATGTAACATTGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28788 ACATAGCAATCTTTAATCAATGTAACATTGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 28847

Qy   1081 TCATCGAGGCCACGCGGGAGTACGATCGAGGGTACGATCGAGGGTACGATGAATAATGCTAGGGAGAGC 1136
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28848 TCATCGAGGCCACGCGGGAGTACGATCGAGGGTACGATCGAGGGTACGATGAATAATGCTAGGGAGAGC 28903
```

RESULT 6
US-10-839-729-15

```
Qy    961 TTGTGCAGAATGAATTCTCGTAACTAAACAGCACAAGTAGGTTTAGTTAACTTTAATCTC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29466 TTGTGCAGAATGAATTCTCGTAACTAAACAGCACAAGTAGGTTTAGTTAACTTTAATCTC 29525

Qy   1021 ACATAGCAATCTTTAATCAATGTGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29526 ACATAGCAATCTTTAATCAATGTGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 29585

Qy   1081 TCATCGAGGCCACGCGGAGTACGAGGGTACAGTGAATAATGCTAGGGAGAGC 1136
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29586 TCATCGAGGCCACGCGGAGTACGAGGGTACAGTGAATAATGCTAGGGAGAGC 29641
```

RESULT 7
US-10-827-757-1
; Sequence 1, Application US/10827757
; Publication No. US2005004071A1
; GENERAL INFORMATION:
; APPLICANT: Comper, Wayne
; TITLE OF INVENTION: Charged Polysaccharides Resistant To Lysosomal Degradation During
; TITLE OF INVENTION: Kidney Filtration And Renal Passage And Their Use To Treat Or
; TITLE OF INVENTION: Prevent Infection By Coronaviruses
; FILE REFERENCE: 11213-007-999
; CURRENT APPLICATION NUMBER: US/10/827,757
; CURRENT FILING DATE: 2004-04-20
; PRIOR APPLICATION NUMBER: 60/464,294
; PRIOR FILING DATE: 2003-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 29727
; TYPE: DNA
; ORGANISM: SARS-related coronavirus (Urbani strain)
US-10-827-757-1

Query Match   100.0%; Score 1136; DB 18; Length 29727;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1136;   Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy      1 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACCATTGGCAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28506 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACCATTGGCAC 28565

Qy     61 CCGCCAATCCTAATAACAATGCTGCCACCGTGCCACCTGCTACAACTTCTCAAGGAACACAACATTGCC 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28566 CCGCCAATCCTAATAACAATGCTGCCACCGTGCCACCTGCTACAACTTCTCAAGGAACACAACATTGCC 28625

Qy    121 AAAAGGCTTCTACCGCAGAGGGAAGCAGAGGCGGCAGTCCGCTCTTCTTCCGCTCCTCATC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28626 AAAAGGCTTCTACCGCAGAGGGAAGCAGAGGCGGCAGTCCGCTCTTCTTCCGCTCCTCATC 28685

Qy    181 ACGTAGTGCGGTAATTCAAGAAATTCAACTCCTGCGCAGCAGTAGGGGAAATTCTCCTGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28686 ACGTAGTGCGGTAATTCAAGAAATTCAACTCCTGCGCAGCAGTAGGGGAAATTCTCCTGC 28745

Qy    241 TCGAATGGCTAGCCGGAGTGGTGAAACTGCCCTCGCCCTTCGGGTAAAGGCCAACAACAAGGCCAAACTGTCACTAA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28746 TCGAATGGCTAGCCGGAGTGGTGAAACTGCCCTCGCCCTTCGGGTAAAGGCCAACAACAAGGCCAAACTGTCACTAA 28805

Qy    301 CCAGCTTGAGAGCAAAGTTTCTGGTAAAGGCCAACAACAACAAGGCCAAACTGTCACTAA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28806 CCAGCTTGAGAGCAAAGTTTCTGGTAAAGGCCAACAACAACAAGGCCAAACTGTCACTAA 28865

Qy    361 GAAATCTGCTGCTGAGGCATCTAAAAAGCCTCGGCCAAAAACGTACTGCCACAAAACAGTA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28866 GAAATCTGCTGCTGAGGCATCTAAAAAGCCTCGGCCAAAAACGTACTGCCACAAAACAGTA 28925

Qy    421 CAACGTCACTCAACTCAAGCAGCATTTGGGACGACGTGGTCAGAGAACAAACCCAAGGAAATTTCGGGGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28926 CAACGTCACTCAACTCAAGCAGCATTTGGGACGACGTGGTCAGAGAACAAACCCAAGGAAATTTCGGGGA 28985

Qy    481 CCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCCAAATTGCACAATTGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  28986 CCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCCAAATTGCACAATTTGC 29045

Qy    541 TCCAAGTGCCTCTGCATCTTCTTGGAATGTCACGCATGGCATTGGCATTGGCATTGGCATTGGCAAATTGCACACCTTCGGG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29046 TCCAAGTGCCTCTGCATCTTCTTGGAATGTCACGCATGGCATTGGCATTGGCATTGGCAAATTGCACACCTTCGGG 29105

Qy    601 AACATGGCTGACTTATCATGGAGCCATTAAATTGATGACAAAGATCCAAAGATCCAAAGATCCAATTCAAAGA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29106 AACATGGCTGACTTATCATGGAGCCATTAAATTGATGACAAAGATCCAAAGATCCAACAATTCAAAGA 29165

Qy    661 CAACGTCATCATACTGCTGAACAAGCACATTGACGAGCTCAGCCTTGCGCGAGAGACAAAAGAACAGAGAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29166 CAACGTCATCATACTGCTGAACAAGCACATTGACGAGCTCAGCCTTGCGCGAGAGACAAAAGAACAGAGAGC 29225

Qy    721 TAAAAAGGACAAAAAGAAAGACTGATGAGCTCAGCCTTGCGCCGCAGAGACAAAAGAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29226 TAAAAAGGACAAAAAGAAAGACTGATGAGCTCAGCCTTGCGCCGCAGAGACAAAAGAA 29285

Qy    781 GCAGCCCACTGTGACTCTTCTTCCTCGCGGCTGACATGGATGATTTCTCCAGACAACTTCA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29286 GCAGCCCACTGTGACTCTTCTTCCTCGCGGCTGACATGGATGATTTCTCCAGACAACTTCA 29345

Qy    841 AAATTCCATGAGTCGGAGCTTCTGCTGATTCAACTCAGGCATAAACACTCATGATGACCAC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29346 AAATTCCATGAGTCGGAGCTTCTGCTGATTCAACTCAGGCATAAACACTCATGATGACCAC 29405

Qy    901 ACAAGGCAGATGGGCTATGTAAAGTTTTCGCAATTCCGTTACGATACATAGTCTACTC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29406 ACAAGGCAGATGGGCTATGTAAAGTTTTCGCAATTCCGTTACGATACATAGTCTACTC 29465

Qy    961 TTGTGCAGAATGAATTCTCGTAACTAAACAGCACAAGTAGGTTTAGTTAACTTTAATCTC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29466 TTGTGCAGAATGAATTCTCGTAACTAAACAGCACAAGTAGGTTTAGTTAACTTTAATCTC 29525

Qy   1021 ACATAGCAATCTTTAATCAATGTGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29526 ACATAGCAATCTTTAATCAATGTGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 29585

Qy   1081 TCATCGAGGCCACGCGGAGTACGAGGGTACAGTGAATAATGCTAGGGAGAGC 1136
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29586 TCATCGAGGCCACGCGGAGTACGAGGGTACAGTGAATAATGCTAGGGAGAGC 29641
```

RESULT 8
US-10-889-447-8
; Sequence 8, Application US/10889447
; Publication No. US2005007530A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 29727
; TYPE: DNA
; ORGANISM: SARS coronavirus Urbani
US-10-889-447-8

Query Match   100.0%; Score 1136; DB 19; Length 29727;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1136;   Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy      1 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACCATTGGCAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28506 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACCATTGGCAC 28565
```

RESULT 9
US-10-699-936-1
; Sequence 1, Application US/10699936
; Publication No. US20050095582A1
; GENERAL INFORMATION:
; APPLICANT: Gillim-Ross, Laura
; APPLICANT: Taylor, Jill
; APPLICANT: Scholl, David R.
; APPLICANT: Wentworth, David E.
; APPLICANT: Jollick, Joseph D.
; TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory
; TITLE OF INVENTION: Syndrome Coronavirus
; FILE REFERENCE: DHI-07986
; CURRENT APPLICATION NUMBER: US/10/699,936
; CURRENT FILING DATE: 2003-11-03
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 29727
; TYPE: DNA
; ORGANISM: SARS coronavirus Urbani
US-10-699-936-1

Query Match   100.0%;   Score 1136;   DB 19;   Length 29727;
Best Local Similarity 100.0%;   Pred. No. 0;
Matches 1136;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

```
Qy      1  AGGCATCGTATGGTTGCAACTGACGGAGCCTTGAACTAACACCCAAAGACACCACATTGGCAC   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28506  AGGCATCGTATGGTTGCAACTGACGGAGCCTTGAACTAACACCCAAAGACACCACATTGGCAC  28565

Qy     61  CCGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGCAACAACATTGCC  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28566  CCGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGCAACAACATTGCC  28625

Qy    121  AAAAGGCTTCTACGCAGAGGAGCAGTAATTCAACTCCTGCTAGACAGATTGAA  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28626  AAAAGGCTTCTACGCAGAGGAGCAGTAATTCAACTCCTGCTAGACAGATTGAA  28685

Qy    181  ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGCTAGGGGAAATTCTCCTGC  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28686  ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGCTAGGGGAAATTCTCCTGC  28745

Qy    241  TCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTGCGGCTATTGCTAGACAGATTGAA  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28746  TCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTGCGGCTATTGCTAGACAGATTGAA  28805

Qy    301  CCAGCTTGAGAGCAAAGTTTCTGGTAAAGGCCAAACTGTCACTAA  360
           |||||||||||||||||||||||||||||||||||||||||||||
Db  28806  CCAGCTTGAGAGCAAAGTTTCTGGTAAAGGCCAAACTGTCACTAA  28865

Qy    361  GAAATCTGCTGCTGAGGTGGTGCCCTGCGGCTATTGCCACAAAACAGTA  420
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db  28866  GAAATCTGCTGCTGAGGTGGTGCCCTGCGGCTATTGCCACAAAACAGTA  28925

Qy    421  CAACGTCACTCAAGCATTGGGAGACGTGGTCCAGAACAAACCCAAGGAAATTTCGGGGA  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28926  CAACGTCACTCAAGCATTGGGAGACGTGGTCCAGAACAAACCCAAGGAAATTTCGGGGA  28985

Qy    481  CCAAGACCTAATCAGACAAGGAACTGATTACAAAACATTGGCCGCAAATTGC  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28986  CCAAGACCTAATCAGACAAGGAACTGATTACAAAACATTGGCCGCAAATTGC  29045

Qy    541  TCCAAGTGCCTCTGAGGCATCTCAAGTCTTTGGAATGTCACGCGTACACCCTTCGGG  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29046  TCCAAGTGCCTCTGAGGCATCTCAAGTCTTTGGAATGTCACGCGTACACCCTTCGGG  29105

Qy    601  AACATGGCTGACTTATCATGGAGCCATTAAATTGGATGACAAAGATCCACAATTCAAGA  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29106  AACATGGCTGACTTATCATGGAGCCATTAAATTGGATGACAAAGATCCACAATTCAAAGA  29165

Qy    661  CAACGTCATACTGCTGAACAAGCAGCACACATTCCCACCACAAAAACAGTA  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db 29166 CAACGTCATACTGCTGAACAAGCACGCACCATTCCCACCGAGAGACAGAGCC 29225
Qy   721 TAAAAAGGACAAAAAGAAAAAGAGACTGATGAAGCTCAGCCTTTGCCGCAGAGACAAAGAA 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29226 TAAAAAGGACAAAAAGAAAAAGAGACTGATGAAGCTCAGCCTTTGCCGCAGAGACAAAGAA 29285
Qy   781 GCAGCCCACTGTGACTCTTCTTCGCGGCTGACATGGATGATTTCTCCAGACAACTTCA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29286 GCAGCCCACTGTGACTCTTCTTCGCGGCTGACATGGATGATTTCTCCAGACAACTTCA 29345
Qy   841 AAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAGGCATAAACACTCATGATGACCAC 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29346 AAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAGGCATAAACACTCATGATGACCAC 29405
Qy   901 ACAAGGCAGATGGGCTATGTAAACGTTTTCGCAATTCCGTTTACGATACATAGTCTACTC 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29406 ACAAGGCAGATGGGCTATGTAAACGTTTTCGCAATTCCGTTTACGATACATAGTCTACTC 29465
Qy   961 TTGTGCAGAATGAATTCTTTAATCAATGTGTAACATTAGTAGTTTAGTTAACTTTAATCTC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29466 TTGTGCAGAATGAATTCTTTAATCAATGTGTAACATTAGTAGTTTAGTTAACTTTAATCTC 29525
Qy  1021 ACATAGCAATCTTTAATCAATGTGTAACATTAGTAGTTTAGTTAACTTTAATCTC 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29526 ACATAGCAATCTTTAATCAATGTGTAACATTAGTAGTTTAGTTAACTTTAATCTC 29585
Qy  1081 TCATCGAGGCCACGCGGAGTCAGTGAATAATGCTAGGGAGAGC 1136
         |||||||||||||||||||||||||||||||||||||||||||
Db 29586 TCATCGAGGCCACGCGGAGTCAGTGAATAATGCTAGGGAGAGC 29641
```

```
Query Match             100.0%;  Score 1136;  DB 18;  Length 29736;
Best Local Similarity   100.0%;
Matches  1136;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     1 AGGCATCGTATGGGTTGCAACTGAGGAGCCTTGAATACACCCAAAGACCACCACATTGGCAC 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28491 AGGCATCGTATGGGTTGCAACTGAGGAGCCTTGAATACACCCAAAGACCACCACATTGGCAC 28550
Qy    61 CCGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGGAACAACATTGCC 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28551 CCGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGGAACAACATTGCC 28610
Qy   121 AAAAGGCTTCTACCGCAGAGGGAAGCAGAGGCGGCAGTCAGGCAGCCTTCTTCGCTCCTCATC 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28611 AAAAGGCTTCTACCGCAGAGGGAAGCAGAGGCGGCAGTCAGGCAGCCTTCTTCGCTCCTCATC 28670
Qy   181 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGCGCAGTAGGGGGAAATTCCTGC 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28671 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGCGCAGTAGGGGGAAATTCCTGC 28730
Qy   241 TCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTCCGCGCTATTGCTGCTAGACAGATTGAA 300
```

```
Db 28731 TCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTCGCGCTATTGCTGCTAGACAGATTGAA 28790
Qy   301 CCAGCTTGAGAGCAAAGTTTCGTGTAAAGGCCAACAACAAGCCAACAACTGTCACTAA 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28791 CCAGCTTGAGAGCAAAGTTTCGTGTAAAGGCCAACAACAAGCCAACAACTGTCACTAA 28850
Qy   361 GAAATCTGCTGCTGAGGCATCTAAAAAGCCTCGCCAAAAAAGGTACTGCCACAAACAGTA 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28851 GAAATCTGCTGCTGAGGCATCTAAAAAGCCTCGCCAAAAAAGGTACTGCCACAAACAGTA 28910
Qy   421 CAACGTCACTCAAGCACTTGGGAGACGTGGTCGTCAGAACAAACCCAAGGAAAATTCGGGGA 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28911 CAACGTCACTCAAGCACTTGGGAGACGTGGTCGTCAGAACAAACCCAAGGAAAATTCGGGGA 28970
Qy   481 CCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAATTGCACAATTTGC 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28971 CCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAATTGCACAATTTGC 29030
Qy   541 TCCAAGTGCCTCTGCATTCTTGGAATGTCACGCATGGAAGTCACACCTTCGGG 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29031 TCCAAGTGCCTCTGCATTCTTGGAATGTCACGCATGGAAGTCACACCTTCGGG 29090
Qy   601 AACATGGCTGACTTATCATGAGCAGCCATTAAATTGGATGACAAAGAAGATCCACAATTCAAAGA 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29091 AACATGGCTGACTTATCATGAGCAGCCATTAAATTGGATGACAAAGAAGATCCACAATTCAAAGA 29150
Qy   661 CAACGTCATACTGCTGAACAAGCACCATACAAAAACATTCCCACCAACAGAGCC 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29151 CAACGTCATACTGCTGAACAAGCACCATACAAAAACATTCCCACCAACAGAGCC 29210
Qy   721 TAAAAAGGACAAAAAGAAAAAGAGACTGATGAAGCTCAGCCTTTGCCGCAGAGACAAAGAA 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29211 TAAAAAGGACAAAAAGAAAAAGAGACTGATGAAGCTCAGCCTTTGCCGCAGAGACAAAGAA 29270
Qy   781 GCAGCCCACTGTGACTCTTCTTCGCGGCTGACATGGATGATTTCTCCAGACAACTTCA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29271 GCAGCCCACTGTGACTCTTCTTCGCGGCTGACATGGATGATTTCTCCAGACAACTTCA 29330
Qy   841 AAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAGGCATAAACACTCATGATGACCAC 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29331 AAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAGGCATAAACACTCATGATGACCAC 29390
Qy   901 ACAAGGCAGATGGGCTATGTAAACGTTTTCGCAATTCCGTTTACGATACATAGTCTACTC 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29391 ACAAGGCAGATGGGCTATGTAAACGTTTTCGCAATTCCGTTTACGATACATAGTCTACTC 29450
Qy   961 TTGTGCAGAATGAATTCTTTAATCAATGTGTAACATTAGTAGTTTAGTTAACTTTAATCTC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29451 TTGTGCAGAATGAATTCTTTAATCAATGTGTAACATTAGTAGTTTAGTTAACTTTAATCTC 29510
Qy  1021 ACATAGCAATCTTTAATCAATGTGTAACATTAGTAGTTTAGTTAACTTTAATCTC 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29511 ACATAGCAATCTTTAATCAATGTGTAACATTAGTAGTTTAGTTAACTTTAATCTC 29570
Qy  1081 TCATCGAGGCCACGCGGAGTCAGTGAATAATGCTAGGGAGAGC 1136
         |||||||||||||||||||||||||||||||||||||||||||
Db 29571 TCATCGAGGCCACGCGGAGTCAGTGAATAATGCTAGGGAGAGC 29626
```

```
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 29736
; TYPE: DNA
; ORGANISM: SARS coronavirus CUHK-W1
US-10-889-447-9

Query Match         100.0%;  Score 1136;  DB 19;  Length 29736;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1136;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACATTGGCAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28491 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACATTGGCAC 28550

Qy     61 CCGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGGCCTCTTCGCTCATC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28551 CCGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGGCCTCTTCGCTCATC 28610

Qy    121 AAAAGGCTTCTACGCAGAGGAAGCAGAGGGCGCAGTCAACACAACAAGGCCAACATTGCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28611 AAAAGGCTTCTACGCAGAGGAAGCAGAGGGCGCAGTCAACACAACAAGGCCAACATTGCC 28670

Qy    181 ACGTAGTCGCGGTAATTCAACATTCAACTTCCTGCAGCAGTAGGGGAAATTCTCCTGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28671 ACGTAGTCGCGGTAATTCAACATTCAACTTCCTGCAGCAGTAGGGGAAATTCTCCTGC 28730

Qy    241 TCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTCCGGCCTATTGCTGCTAGACAGATTGAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28731 TCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTCCGGCCTATTGCTGCTAGACAGATTGAA 28790

Qy    301 CCAGCTTGAGAGCAAAGTTTCTGGTAAAGGCCAACAAGGCCAAACTGTCACTAA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28791 CCAGCTTGAGAGCAAAGTTTCTGGTAAAGGCCAACAAGGCCAAACTGTCACTAA 28850

Qy    361 GAAATCTGCTGCTGAGGCATCTAAAAAGCCTCGCCAAAAACGTACTGCCAAAAACAGTA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28851 GAAATCTGCTGCTGAGGCATCTAAAAAGCCTCGCCAAAAACGTACTGCCAAAAACAGTA 28910

Qy    421 CAACGTCACTCAAGCACTTGGGAGACGTGGTCCAGAACAAACCCAAGGAAATTTCGGGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28911 CAACGTCACTCAAGCACTTGGGAGACGTGGTCCAGAACAAACCCAAGGAAATTTCGGGA 28970

Qy    481 CCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAATTGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28971 CCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAATTGC 29030

Qy    541 TCCAAGTGCCCTCGCAATTCTTGGAATGTCACGCATTGGCATGGAAGTCACACCTTCGGG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29031 TCCAAGTGCCCTCGCAATTCTTGGAATGTCACGCATTGGCATGGAAGTCACACCTTCGGG 29090

Qy    601 AACATGGCTGACTTATCATGAGCCATTAAATTGGATGACAAAGATCCACAATTCAAAGA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29091 AACATGGCTGACTTATCATGAGCCATTAAATTGGATGACAAAGATCCACAATTCAAAGA 29150

Qy    661 CAACGTCATCACTGCTGAACAAGCCACATCCCGCGCAGACAAAAGATTCCCACCAACAGAGCC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29151 CAACGTCATCACTGCTGAACAAGCCACATCCCGCGCAGACAAAAGATTCCCACCAACAGAGCC 29210

Qy    721 TAAAAAGGACAAAAAGAAAAAAGACTGATGAAGCTCAGCCTTTGCCGCCGAGACAAAAGA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29211 TAAAAAGGACAAAAAGAAAAAAGACTGATGAAGCTCAGCCTTTGCCGCCGAGACAAAAGA 29270

Qy    781 GCAGCCCACTGTGACTCTTCCTCCTCCGCCTGACACTGACGATTCTTCCAGACAACTTCA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29271 GCAGCCCACTGTGACTCTTCCTCCTCCGCCTGACACTGACGATTCTTCCAGACAACTTCA 29330

Qy    841 AAATTCCATGAGTGAGCTTGCTCGTGATTCTCAGGCATCAAACACTCATGATGACCAC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29331 AAATTCCATGAGTGAGCTTGCTCGTGATTCTCAGGCATCAAACACTCATGATGACCAC 29390

Qy    901 ACAAGGCAGATGGGCTATGTGAACGTTTCCGTTTACGATACATAGTCTACTC
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29391 ACAAGGCAGATGGGCTATGTGAACGTTTCCGTTTACGATACATAGTCTACTC 29450
```

```
Qy    961 TTGTGCAGAATGAATTCTCGTAACTAAACAGACAAGTAGGTTAGTTAACTTTAATCTC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29451 TTGTGCAGAATGAATTCTCGTAACTAAACAGACAAGTAGGTTAGTTAACTTTAATCTC 29510

Qy   1021 ACATAGCAATCTTAATCAATGTAACATTAGGGAGGACTTGAAAGAGGCCACCACATT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29511 ACATAGCAATCTTAATCAATGTAACATTAGGGAGGACTTGAAAGAGGCCACCACATT 29570

Qy   1081 TCATCGAGGCCACGCGCGGAGTACGATCGAGGGTACAGTGAATATGCTAGGGAGAGC 1136
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29571 TCATCGAGGCCACGCGCGGAGTACGATCGAGGGTACAGTGAATATGCTAGGGAGAGC 29626
```

```
RESULT 12
US-10-699-936-3
; Sequence 3, Application US/1069936
; Publication No. US2005009558281
; GENERAL INFORMATION:
;   APPLICANT: Gillim-Ross, Laura
;   APPLICANT: Taylor, Jill
;   APPLICANT: Scholl, David R.
;   APPLICANT: Wentworth, David E.
;   APPLICANT: Jollick, Joseph D.
;   TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory
;   TITLE OF INVENTION: Syndrome Coronavirus
;   FILE REFERENCE: DHI-07986
;   CURRENT APPLICATION NUMBER: US/10/699,936
;   CURRENT FILING DATE: 2003-11-03
;   NUMBER OF SEQ ID NOS: 87
;   SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
;   LENGTH: 29736
;   TYPE: DNA
;   ORGANISM: SARS coronavirus CUHK-W1
US-10-699-936-3

Query Match         100.0%;  Score 1136;  DB 19;  Length 29736;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1136;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACATTGGCAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28491 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACATTGGCAC 28550

Qy     61 CCGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGGAACAACATTGCC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28551 CCGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGGAACAACATTGCC 28610

Qy    121 AAAAGGCTTCTACGCAGAGGAAGCAGAGGGCGCAGTCAACACAACAAGGCCAACATTGCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28611 AAAAGGCTTCTACGCAGAGGAAGCAGAGGGCGCAGTCAACACAACAAGGCCAACATTGCC 28670

Qy    181 ACGTAGTCGCGGTAATTCAACATTCAACTTCCTGCAGCAGTAGGGGAAATTCTCCTGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28671 ACGTAGTCGCGGTAATTCAACATTCAACTTCCTGCAGCAGTAGGGGAAATTCTCCTGC 28730

Qy    241 TCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTCCGGCCTATTGCTGCTAGACAGATTGAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28731 TCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTCCGGCCTATTGCTGCTAGACAGATTGAA 28790

Qy    301 CCAGCTTGAGAGCAAAGTTTCTGGTAAAGGCCAACAAGGCCAAACTGTCACTAA 350
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28791 CCAGCTTGAGAGCAAAGTTTCTGGTAAAGGCCAACAAGGCCAAACTGTCACTAA 28850

Qy    361 GAAATCTGCTGCTGAGGCATCTAAAAAGCCTCGCCAAAAACGTACTGCCAAAAACAGTA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28851 GAAATCTGCTGCTGAGGCATCTAAAAAGCCTCGCCAAAAACGTACTGCCAAAAACAGTA 28910

Qy    421 CAACGTCACTCAAGCACTTGGGAGACGTGGTCCAGAACAAACCCAAGGAAATTTCGGGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28911 CAACGTCACTCAAGCACTTGGGAGACGTGGTCCAGAACAAACCCAAGGAAATTTCGGGA 28970
```

```
Db 28506 AGGCATCGTATGGGTTGCAACTGAGGAGCCTTGAATACACCCAAAGACCACATTGGCAC 28565
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 28911 CAACGTCACTCAAGCATTGGGAGACGTGGTCCAGAACAAACCCAAGAAATTGCACAATTGC 28970

Qy    61 CCGCAATCCTAATAACAATGCTCGCCACCGTGCTACAACTTCCTCAGGAACAACATTGCC 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28566 CCGCAATCCTAATAACAATGCTCGCCACCGTGCTACAACTTCCTCAGGAACAACATTGCC 28625

Qy   121 AAAAGGCTTCTACCGAGGGAAGCAGAGGCGGCAGTCACCTTCGCCTCTTCTCGCTCCTCATC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28626 AAAAGGCTTCTACCGAGGGAAGCAGAGGCGGCAGTCACCTTCGCCTCTTCTCGCTCCTCATC 28685

Qy   181 ACGTAGTCGCGTAATTCAAGAAATTCAACTCCTGGCAGCAGTAGGGGAAATTCTCCTGC 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28686 ACGTAGTCGCGTAATTCAAGAAATTCAACTCCTGGCAGCAGTAGGGGAAATTCTCCTGC 28745

Qy   241 TCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTCGCGTATTGCTGTCTAGACAGATTGAA 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28746 TCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTCGCGTATTGCTGTCTAGACAGATTGAA 28805

Qy   301 CCAGCTTGAGAGACAAAGTTTCTGTGTAAAGGCCAACAACAACAAGGCCAAACTGTCACTAA 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28806 CCAGCTTGAGAGACAAAGTTTCTGTGTAAAGGCCAACAACAACAAGGCCAAACTGTCACTAA 28865

Qy   361 GAAATCTGCTGCTGAGGCATCTAAAAAGCTCGGCCAAAAACGTACTGCCACAAAACAGTA 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28866 GAAATCTGCTGCTGAGGCATCTAAAAAGCTCGGCCAAAAACGTACTGCCACAAAACAGTA 28925

Qy   421 CAACGTCACTCAAGCATTTGGGAGACGTGGTCCAGAACAAACCCAAGGAAATTTCGGGGA 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28926 CAACGTCACTCAAGCATTTGGGAGACGTGGTCCAGAACAAACCCAAGGAAATTTCGGGGA 28985

Qy   481 CCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAATTGCACAATTTGC 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28986 CCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAATTGCACAATTTGC 29045

Qy   541 TCCAAGTGCCTCTGCCATTCTTGGAATGTCACGCATTGGAAGTGAAGTCACACCTTCGGG 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29046 TCCAAGTGCCTCTGCCATTCTTGGAATGTCACGCATTGGAAGTGAAGTCACACCTTCGGG 29105

Qy   601 AACATGGCTGACTTATCATGAGCCATAAATTGGATGACAAAGATCCACAATTCAAAGA 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29106 AACATGGCTGACTTATCATGAGCCATAAATTGGATGACAAAGATCCACAATTCAAAGA 29165

Qy   661 CAAGTCATCATACTGCTGAACAAGCACATTGACGACATACAAAACATTCCCACACCAGAGCC 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29166 CAAGTCATCATACTGCTGAACAAGCACATTGACGACATACAAAACATTCCCACACCAGAGCC 29225

Qy   721 TAAAAAGGACAAAAAGAAAAAGACTGATGAAGCTGATGAAGCTCAGCCTTTGCGCGAGAGACAAAAGAA 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29226 TAAAAAGGACAAAAAGAAAAAGACTGATGAAGCTGATGAAGCTCAGCCTTTGCGCGAGAGACAAAAGAA 29285

Qy   781 GCAGCCACTGACTCTTCTTCCTCGCGGCTGACATGGATGATTTCTCCAGACAACTTCA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29286 GCAGCCACTGACTCTTCTTCCTCGCGGCTGACATGGATGATTTCTCCAGACAACTTCA 29345

Qy   841 AAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAGGCATAAACACTCATGATGACCAC 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29346 AAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAGGCATAAACACTCATGATGACCAC 29405

Qy   901 ACAAGGCAGATGGGCTATGTAAAGTTTTCGCAATTCCGTTTACAAGTTCGTTTAGTTTAGTCTACTC 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29406 ACAAGGCAGATGGGCTATGTAAAGTTTTCGCAATTCCGTTTACAAGTTCGTTTAGTTTAGTCTACTC 29465

Qy   961 TTGTGCAGAATGAATTCTCGTAACTAAACAGCACAAGCACAAGTAGGTTTAGTTTAATCTC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29466 TTGTGCAGAATGAATTCTCGTAACTAAACAGCACAAGCACAAGTAGGTTTAGTTTAATCTC 29525

Qy  1021 ACATAGCAATCTTTAATCAATGTGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29526 ACATAGCAATCTTTAATCAATGTGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 29585

Qy  1081 TCATCGAGCCACCGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC 1136
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29571 TCATCGAGCCACCGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC 29626
```

.Db   29586 TCATCGAGGCCACGCGCGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC 29641

RESULT 14
US-10-808-187-15
; Sequence 15, Application US/10808187
; Publication No. US20050009009A1
; GENERAL INFORMATION:
;   APPLICANT: PEIRIS, JOSEPH S. M.
;   APPLICANT: YUEN, KWOK YUNG
;   APPLICANT: POON, LIT MAN
;   APPLICANT: GUAN, YI
;   APPLICANT: CHAN, KWOK HUNG
;   APPLICANT: NICHOLLS, JOHN
;   TITLE OF INVENTION: A DIAGNOSTIC ASSAY FOR THE HUMAN VIRUS CAUSING SEVERE ACUTE
;   TITLE OF INVENTION: RESPIRATORY SYNDROME (SARS)
;   FILE REFERENCE: V9661.0078
;   CURRENT APPLICATION NUMBER: US/10/808,187
;   CURRENT FILING DATE: 2004-03-24
;   PRIOR APPLICATION NUMBER: 60/457,031
;   PRIOR FILING DATE: 2003-03-24
;   PRIOR APPLICATION NUMBER: 60/457,730
;   PRIOR FILING DATE: 2003-03-26
;   PRIOR APPLICATION NUMBER: 60/459,931
;   PRIOR FILING DATE: 2003-04-02
;   PRIOR APPLICATION NUMBER: 60/460,357
;   PRIOR FILING DATE: 2003-04-03
;   PRIOR APPLICATION NUMBER: 60/461,265
;   PRIOR FILING DATE: 2003-04-08
;   PRIOR APPLICATION NUMBER: 60/462,805
;   PRIOR FILING DATE: 2003-04-14
;   PRIOR APPLICATION NUMBER: 60/468,139
;   PRIOR FILING DATE: 2003-05-05
;   PRIOR APPLICATION NUMBER: 60/464,886
;   PRIOR FILING DATE: 2003-04-23
;   PRIOR APPLICATION NUMBER: 60/471,200
;   PRIOR FILING DATE: 2003-05-16
;   NUMBER OF SEQ ID NOS: 2476
;   SOFTWARE: PatentIn ver. 3.2
; SEQ ID NO 15
;   LENGTH: 29742
;   TYPE: DNA
;   ORGANISM: Human severe acute respiratory system virus
US-10-808-187-15

Query Match   100.0%; Score 1136; DB 19; Length 29742;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1136; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACCATTGGCAC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28506 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACCATTGGCAC 28565

Qy   61 CCGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGGAACAACATTGCC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28566 CCGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGGAACAACATTGCC 28625

Qy   121 AAAAGGCTTCTACCGCAGAGGGAAGCAGAGGCGGCAGTCTTTCTTCGCGTCCTCCTCATC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28626 AAAAGGCTTCTACCGCAGAGGGAAGCAGAGGCGGCAGTCTTTCTTCGCGTCCTCCTCATC 28685

Qy   181 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGCGCAGCAGTAGGGGAAATTCCTGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28686 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGCGCAGCAGTAGGGGAAATTCCTGC 28745

Qy   241 TCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTCCGCGCTATTGCTGCTAGCACAGATTGAA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28746 TCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTCCGCGCTATTGCTGCTAGCACAGATTGAA 28805

Qy   301 CCAGCTTGAGAGCAAAGTTTCTGGTAAAGGCCAAACAACAGGCCAAACTGTCACTAA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28806 CCAGCTTGAGAGCAAAGTTTCTGGTAAAGGCCAAACAACAGGCCAAACTGTCACTAA 28865

Qy   361 GAAATCTGCTGCTGAGGCATCTCAAAAAGCCTCGCCAAAAACTACTGCCACAAAACAGTA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28866 GAAATCTGCTGCTGAGGCATCTCAAAAAGCCTCGCCAAAAACTACTGCCACAAAACAGTA 28925

Qy   421 CAACGTCACTCAAGCATTGGGAGACGTGGTCAGAGACAAACCCAAGGAAATTTCGGGGA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28926 CAACGTCACTCAAGCATTGGGAGACGTGGTCAGAGACAAACCCAAGGAAATTTCGGGGA 28985

Qy   481 CCAAGACCTAATCAGACAGGAACTGATTACAACATTGGCCGCAAATTGCACAATTTGC 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 28986 CCAAGACCTAATCAGACAGGAACTGATTACAACATTGGCCGCAAATTGCACAATTTGC 29045

Qy   541 TCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCACACCTTCGGG 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29046 TCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCACACCTTCGGG 29105

Qy   601 AACATGGCTGACTTATCATGCCATTAAATTGGATGACAAAGATCACACATTCAAAGA 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29106 AACATGGCTGACTTATCATGCCATTAAATTGGATGACAAAGATCACACATTCAAAGA 29165

Qy   661 CAACGTCATACTGCTGACAAGCACATTCCCACCAACAGAGCC 720
           ||||||||||||||||||||||||||||||||||||||||||
Db 29166 CAACGTCATACTGCTGACAAGCACATTCCCACCAACAGAGCC 29225

Qy   721 TAAAAAGGACAAAAAGAAAAGACTGATGAGCTCAGCCTTTGCCCGAGAGACAAAAGAA 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29226 TAAAAAGGACAAAAAGAAAAGACTGATGAGCTCAGCCTTTGCCCGAGAGACAAAAGAA 29285

Qy   781 GCAGCCCACTGACTCTTCCTCCTGCGGGCTGACATGGATGATTTCCAGACAACTTCA 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29286 GCAGCCCACTGACTCTTCCTCCTGCGGGCTGACATGGATGATTTCCAGACAACTTCA 29345

Qy   841 AAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAGGCATAAACACTCATGATGACCAC 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29346 AAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAGGCATAAACACTCATGATGACCAC 29405

Qy   901 ACAAGGCCAGATGGGCTATGTAAAAGTTTCGCAATTCCGTTTACGATACATAGTCTACTC 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29406 ACAAGGCCAGATGGGCTATGTAAAAGTTTCGCAATTCCGTTTACGATACATAGTCTACTC 29465

Qy   961 TTGTGCAGAATGAATTCTCGTAACTAAACAGCACAAGAGTAGGTTTAGTTAACTTAATCTC 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29466 TTGTGCAGAATGAATTCTCGTAACTAAACAGCACAAGAGTAGGTTTAGTTAACTTAATCTC 29525

Qy   1021 ACATAGCAATCTTTAATCAATGTGTAACATGTGAAAGAGGACTTGAAAGAGCCACCACATTT 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29526 ACATAGCAATCTTTAATCAATGTGTAACATGTGAAAGAGGACTTGAAAGAGCCACCACATTT 29585

Qy   1081 TCATCGAGGCCACGCGCGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC 1136
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 29586 TCATCGAGGCCACGCGCGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC 29641

RESULT 15
US-10-808-187-16
; Sequence 16, Application US/10808187
; Publication No. US20050009009A1
; GENERAL INFORMATION:
;   APPLICANT: PEIRIS, JOSEPH S. M.
;   APPLICANT: YUEN, KWOK YUNG
;   APPLICANT: POON, LIT MAN
;   APPLICANT: GUAN, YI
;   APPLICANT: CHAN, KWOK HUNG
;   APPLICANT: NICHOLLS, JOHN
;   TITLE OF INVENTION: A DIAGNOSTIC ASSAY FOR THE HUMAN VIRUS CAUSING SEVERE ACUTE
;   TITLE OF INVENTION: RESPIRATORY SYNDROME (SARS)
;   FILE REFERENCE: V9661.0078
;   CURRENT APPLICATION NUMBER: US/10/808,187
;   CURRENT FILING DATE: 2004-03-24
;   PRIOR APPLICATION NUMBER: 60/457,031
;   PRIOR FILING DATE: 2003-03-24
;   PRIOR APPLICATION NUMBER: 60/457,730
;   PRIOR FILING DATE: 2003-03-26

```
;   PRIOR APPLICATION NUMBER: 60/459,931
;   PRIOR FILING DATE: 2003-04-02
;   PRIOR APPLICATION NUMBER: 60/460,357
;   PRIOR FILING DATE: 2003-04-03
;   PRIOR APPLICATION NUMBER: 60/461,265
;   PRIOR FILING DATE: 2003-04-08
;   PRIOR APPLICATION NUMBER: 60/462,805
;   PRIOR FILING DATE: 2003-04-14
;   PRIOR APPLICATION NUMBER: 60/468,139
;   PRIOR FILING DATE: 2003-05-05
;   PRIOR APPLICATION NUMBER: 60/464,886
;   PRIOR FILING DATE: 2003-04-23
;   PRIOR APPLICATION NUMBER: 60/471,200
;   PRIOR FILING DATE: 2003-05-16
;   NUMBER OF SEQ ID NOS: 2476
;   SOFTWARE: PatentIn ver. 3.2

;   SEQ ID NO 16
;   LENGTH: 29742
;   TYPE: DNA
;   ORGANISM: Human severe acute respiratory system virus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(29742)
US-10-808-187-16
```

```
Query Match           100.0%;  Score 1136;  DB 19;  Length 29742;
Best Local Similarity 100.0%;  Pred. No.0;
Matches 1136;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      1 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATAACACCCAAAGACCACATTGGCAC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28506 AGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATAACACCCAAAGACCACATTGGCAC 28565

Qy     61 CCGCAATCCTAATAACAATGCTGCCACCGTGCTACACACTTCCTCAAGGAACAACATTGCC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28566 CCGCAATCCTAATAACAATGCTGCCACCGTGCTACACACTTCCTCAAGGAACAACATTGCC 28625

Qy    121 AAAAGGCTTCTACGCAGAGGAAGCAGAGGCGGCAGTCGAACTCCTTCGCTCCTCCTCATC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28626 AAAAGGCTTCTACGCAGAGGAAGCAGAGGCGGCAGTCGAACTCCTTCGCTCCTCCTCATC 28685

Qy    181 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGGCAGCAGTAGGGGAAATTCTCCTGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28686 ACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGGCAGCAGTAGGGGAAATTCTCCTGC 28745

Qy    241 TCGAATGGCTAGCCGGAGGTGGTGAAACTGCCCTCGCCTATTGCTAGACAGATTGAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28746 TCGAATGGCTAGCCGGAGGTGGTGAAACTGCCCTCGCCTATTGCTAGACAGATTGAA 28805

Qy    301 CCAGCTTGAGAGAGCAAAGTTTCTGTGAAAGGCCAACAACAACAAGGCCAAACTGTCACTAA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28806 CCAGCTTGAGAGAGCAAAGTTTCTGTGAAAGGCCAACAACAACAAGGCCAAACTGTCACTAA 28865

Qy    361 GAAATCTGCTGCTGAGGCATCTAAAAAAGCCTAAAAAAGCCTCGCCAAAAACGTACTACTGCCACAAAACAGTA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28866 GAAATCTGCTGCTGAGGCATCTAAAAAAGCCTAAAAAAGCCTCGCCAAAAACGTACTACTGCCACAAAACAGTA 28925

Qy    421 CAACGTCACTCAAGCATTTGGGACGCGTGTCCAGAACAAACCCAAGGAAATTTCGGGGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28926 CAACGTCACTCAAGCATTTGGGACGCGTGTCCAGAACAAACCCAAGGAAATTTCGGGGA 28985

Qy    481 CCAAGAACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAATTGCACAATTTGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28986 CCAAGAACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAATTGCACAATTTGC 29045

Qy    541 TCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCACACCTTCGGG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29046 TCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCACACCTTCGGG 29105

Qy    601 AACATGGCTGACTTATCATGGAAGCCATTAAATTGGATGACAAAGATCACCAATTCAAAGA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29106 AACATGGCTGACTTATCATGGAAGCCATTAAATTGGATGACAAAGATCACCAATTCAAAGA 29165

Qy    661 CAACGTCATACTGCTGAACAAGCACACATTGACGCATACAAAACATTCCCACCAACAGAGCC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29166 CAACGTCATACTGCTGAACAAGCACACATTGACGCATACAAAACATTCCCACCAACAGAGCC 29225

Qy    721 TAAAAAGGACAAAAAGAAAAGACTGATGAAGCTTCAGCCTTTGCGCAGAGACAAAAAGAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29226 TAAAAAGGACAAAAAGAAAAGACTGATGAAGCTTCAGCCTTTGCGCAGAGACAAAAAGAA 29285

Qy    781 GCAGCCCACTGTGACTCTTCTTCCTGCGGCTGACATGGATGATTCTCCAGACAACTTCA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29286 GCAGCCCACTGTGACTCTTCTTCCTGCGGCTGACATGGATGATTCTCCAGACAACTTCA 29345

Qy    841 AAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAGGCATAAAACACTCATGATGACCAC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29346 AAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAGGCATAAAACACTCATGATGACCAC 29405

Qy    901 ACAAGGCAGATGGGCTATGTAAAGTTTTCGCAATTCCGTTACGATACATAGTCTACTC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29406 ACAAGGCAGATGGGCTATGTAAAGTTTTCGCAATTCCGTTACGATACATAGTCTACTC 29465

Qy    961 TTGTGCAGAATGAATTCTCGTAACTAAACAGCACAAGTAGGTTTAGTTAACTTTAATCTC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29466 TTGTGCAGAATGAATTCTCGTAACTAAACAGCACAAGTAGGTTTAGTTAACTTTAATCTC 29525

Qy   1021 ACATAGCAATCTTTAATCAATGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29526 ACATAGCAATCTTTAATCAATGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTT 29585

Qy   1081 TCATCGAGGCCACCGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC 1136
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29586 TCATCGAGGCCACCGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGAGC 29641
```

Search completed: May 15, 2005, 16:48:43
Job time : 843.5 secs

```
                  GenCore version 5.1.6
        Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:          May 15, 2005, 11:03:36 ; Search time 5001.54 Seconds
                                           (without alignments)
                 8645.536 Million cell updates/sec

Title:           US-10-764-075-1
Perfect score:   1136
Sequence:        1 aggcaccgtatgggtgcaa.........tgaataatgctaggagagc 1136

Scoring table:   IDENTITY_NUC
                 Gapop 10-0 , Gapext 1.0

Searched:        34239544 seqs, 1903213470 residues

Total number of hits satisfying chosen parameters:    68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       EST:*
                 1: gb_est1:*
                 2: gb_est2:*
                 3: gb_htc:*
                 4: gb_est3:*
                 5: gb_est4:*
                 6: gb_est5:*
                 7: gb_est6:*
                 8: gb_gss1:*
                 9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
c 1	43.2	3.8	769	6	CB669132	CB669132 OSJNEe01G
c 2	43	3.8	323	7	H55033	H55033 HHU58b Sorg
c 3	42.8	3.8	640	4	BJ328127	BJ328127 BJ328127
4	42.4	3.7	1101	9	CNS00LQ0	AL068607 Drosophil
5	42.2	3.7	434	8	BJ358542	BJ358542 BJ358542
6	42.2	3.7	518	2	BF598760	BF598760 sv21d10.y
7	42.2	3.7	540	4	BJ362480	BJ362480 BJ362480
c 8	42.2	3.7	556	4	BJ366773	BJ366773 BJ366773
9	42.2	3.7	593	4	BJ328471	BJ328471 BJ328471
c 10	42.2	3.7	601	4	BJ339144	BJ339144 BJ339144
c 11	42.2	3.7	605	4	BJ365525	BJ365525 BJ365525
c 12	42.2	3.7	606	4	BJ339537	BJ339537 BJ339537
c 13	42.2	3.7	616	4	BJ387757	BJ387757 BJ387757
c 14	42.2	3.7	617	4	BJ366277	BJ366277 BJ366277
c 15	42.2	3.7	629	4	BJ370135	BJ370135 BJ370135
c 16	42.2	3.7	656	4	BJ361324	BJ361324 BJ361324
17	42	3.7	532	5	BX512459	BX512459 BX512459
c 18	42	3.7	583	4	BJ333936	BJ333936 BJ333936
19	42	3.7	632	2	BB623030	BB623030 AGENCOURT
20	42	3.7	772	5	BU558830	BU558830 AGENCOURT
21	42	3.7	780	9	AY400589	AY400589 Mus muscu
22	42	3.7	1002	6	BY708644	BY708644 Mus muscu
23	42	3.7	1411	3	AK008741	AK008741 Mus muscu
24	42	3.7	1532	3	AK017592	AK017592 Mus muscu
c 25	42	3.7	2310	3	AK032624	AK032624 Mus muscu
c 26	42	3.7	2718	3	AK032314	AK032314 Mus muscu
c 27	42	3.7	3658	3	AK083404	AK083404 Mus muscu
c 28	41.6	3.7	536	4	BJ387830	BJ387830 BJ387830
c 29	41.4	3.6	300	8	AZ254800	AZ254800 p050 K19L
30	41.4	3.6	777	9	AG477863	AG477863 Mus muscu
31	41.4	3.6	834	9	CNS0112M	AL099736 Drosophil
32	41.4	3.6	1049	9	CNS00IW5	AL075516 Drosophil
c 33	41.4	3.6	1101	9	CNS00LT2	AL078714 Drosophil
c 34	40.8	3.6	922	9	CNS0073W	AL066784 Drosophil
35	40.6	3.6	467	2	BF286190	BF286190 EST450781
36	40.6	3.6	517	2	BF566518	BF566518 UI-R-BO0-
c 37	40.6	3.6	558	2	AW919227	AW919227 EST350531
38	40.6	3.6	723	2	BH961520	BH961520 odd50g10.
c 39	40.6	3.6	1117	8	CC227680	CC227680 CH261-44M
40	40.4	3.6	463	1	AI060840	AI060840 ub44e06.r
41	40.2	3.6	294	5	BB125565	BB125565 BB125565
c 42	39.8	3.5	353	6	CA780863	CA780863 MPL384_7
43	39.8	3.5	693	4	BG644557	BG644557 EST506I76
c 44	39.8	3.5	747	9	AG375742	AG375742 Mus muscu
45	39.8	3.5	779	9	CNS04A9H	AL281582 Tetraodon

ALIGNMENTS

```
RESULT 1
CB669132/c
LOCUS        CB669132              769 bp    mRNA    linear   EST 09-APR-2003
DEFINITION   OSJNEe01G10.f OSJNEe Oryza sativa (japonica cultivar-group) cDNA
             clone OSJNEe01G10 5', mRNA sequence.
ACCESSION    CB669132
VERSION      CB669132.1  GI:29672857
KEYWORDS     EST.
SOURCE       Oryza sativa (japonica cultivar-group)
ORGANISM     Oryza sativa (japonica cultivar-group)
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
             Ehrhartoideae; Oryzeae; Oryza.
REFERENCE    1  (bases 1 to 769)
AUTHORS      Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
             Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE        Large-scale identification of ESTs involved in the interaction
             between rice and Magnaporthe grisea
JOURNAL      Unpublished (2003)
COMMENT      Contact: Rod Wing
             Arizona Genomics Institute
             University of Arizona
             Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
             85721-0088, USA
             Tel: 520 626 3967
             Fax: 520 621 9288
             Email: http://genome.arizona.edu
             PCR PRimers
             FORWARD: gta aaa cga cgg cca gtg
             BACKWARD: gga aac agc tat gac cat g
             Plate: 01  row: G  column: 10
             Seq primer: gta aaa cga cgg cca gtg.
FEATURES             Location/Qualifiers
     source          1..769
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="OSJNEe01G10"
                     /tissue_type="Leaf"
                     /dev_stage="3 week"
                     /lab_host="DH10B"
                     /clone_lib="OSJNEe"
                     /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                     XhoI; 24 hrs after innoculation with Rice Blast (70-15)"
ORIGIN
```

```
Query Match           3.8%;  Score 43.2;  DB 6;  Length 769;
Best Local Similarity 48.8%;  Pred. No. 0.55;
Matches  117;  Conservative   0;  Mismatches  123;  Indels   0;  Gaps   0;

Qy   524 CAAATTGCACAATTTGCTCCAAGTGCCTCTGCCATTCTTTGGAATGTCACGCATTGGCATG 583
                |||  ||  |  |||||  ||| | ||| | || |||  |||||||| || ||
Db   444 CGACGTGCTATCTTTGAGCCAATTTTTAGTTTTTGCTTTGGGTTGATCCCATACTCGTTT 385

Qy   584 GAAGTCACACCTTCGGGAACATGGCTACTTATCATGGAGCCATTAAATTGGATGACAAA 643
                |  | || |||  |||      ||| ||| |||  |||| |||    |||| ||||
Db   384 GGAATTACCCCATCAGCAAGAGCTACACAATCAATGTTAGATAAGCTATGATGAAAAT 325

Qy   644 GATCCACCAATTCAAAGACAACAACGTCATCTGCTGAACAGACACATTGACGCATCAAAACA 703
                || |||||| |||| ||        ||     |||||  |  | ||||  | || |||||
Db   324 ATTAAAAACGTCAAAGAGAACTAATCTACACAAATAAACTAACAAAAAAATCCAAAACT 265

Qy   704 TTCCCACCACAGAGCCTAAAAGGACAAAAGAAAAAGACTGATGAGAGCTCAGCCCTTTG 763
                ||   ||| |||  |||  |||    | |||       ||||||||||   ||||||
Db   264 ATAACATCATATGTGGCCTCTAATAGATAAAAGTATAATTTTTCAGAGGGTCAAACTTTG 205

RESULT 2
H55033/c
LOCUS       H55033                   323 bp    mRNA    linear   EST 27-SEP-1999
DEFINITION  HHU58b Sorghum bicolor cv. TX430 Sorghum bicolor cDNA clone HHU58
            3' similar to transketolase, chloroplast (TKLC1), mRNA sequence.
ACCESSION   H55033
VERSION     H55033.1  GI:2674237
KEYWORDS    EST.
SOURCE      Sorghum bicolor (sorghum)
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 323)
  AUTHORS   Wyrich,R., Dressen,U., Brockmann,S., Streubel,M., Chang,C.,
            Qiang,D., Paterson.A.H. and Westhoff,P.
  TITLE     The molecular basis of C4 photosynthesis in sorghum: isolation,
            characterization and RFLP mapping of mesophyll- and
            bundle-sheath-specific cDNAs obtained by differential screening
  JOURNAL   Plant Mol. Biol. 37 (2), 319-335 (1998)
  MEDLINE   98278366
   PUBMED   9617804
COMMENT     On Dec 10, 1997 this sequence version replaced gi:1000813.
            Contact: Westhoff P
            Plant Molecular Biology
            Heinrich-Heine-University
            Universitaetsstrasse, D-40225 Duesseldorf, Germany
            Tel: (49)-211-311-2338
            Fax: (49)-211-311-4871
            Email: west@uni-duesseldorf.de
            Homology: sp/p22976, Blastx-Score: 176, pVal: 2.1e-29
            Seq primer: M13 (-20).
FEATURES             Location/Qualifiers
     source          1..323
                     /organism="Sorghum bicolor"
                     /mol_type="mRNA"
                     /strain="cv. TX430"
                     /db_xref="taxon:4558"
                     /clone="HHU58"
                     /clone_lib="Sorghum bicolor cv. TX430"
                     /note="Vector: Lambda ZAP II; leaf cDNA library,
                     unidirectionally cloned"
ORIGIN

Query Match           3.8%;  Score 43;  DB 7;  Length 323;
Best Local Similarity 52.6%;  Pred. No. 0.48;
Matches   91;  Conservative   0;  Mismatches  82;  Indels   0;  Gaps   0;

Qy   627 TTAAATTGGATGACAAAGATCCACCAATTCAAAGACAACAACGTCATCTGCTGAACAGCACA 686
                ||    ||| |   || |   |||| |||  |   |||  | ||   ||
Db   313 TTTTTTTTGAAGAAAATAACATCACATTATTGAGCTCAATTAGTTATGTTATGTACCATA 254
```

```
Qy   687 TTGACGCATACAAAACATTCCACCACAGAGCCTTAAAAGGACAAAAAGAAAAAGAACTG 746
                || |   ||||||  ||       ||    ||||||||||||||||| ||||||||
Db   253 AAAGCCAGTAGAAAACTGNNTCAATAAAAGAGCAAAAAAAAATGCAACAACAATAAGACAC 194

Qy   747 ATGAAGCTCAGCCTTGCCGCAGAGACAAAAGAGGAGCAGCCCACTGTGACTCTT 799
                ||||| ||  |  | || || ||| |||  || || ||||| |  ||| |||
Db   193 GTGCAGCTTAACCCACGATCAGCAAGAGATGAAGCATCGTCCTCTGCAAGACGT 141

RESULT 3
BJ328127/c
LOCUS       BJ328127                 640 bp    mRNA    linear   EST 05-MAR-2002
DEFINITION  BJ328127 Dictyostelium discoideum cDNA library, AF Dictyostelium
            discoideum cDNA clone dda23e04 5', mRNA sequence.
ACCESSION   BJ328127
VERSION     BJ328127.1  GI:19158257
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyostelida; Dictyosteliida.
REFERENCE   1  (bases 1 to 640)
  AUTHORS   Urushihara,H., Tanaka,Y., Kohara,Y., and Shin-i,T.
  TITLE     Full length cDNA of Dictyostelium discoideum at the aggregation
            stage
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1..640
                     /organism="Dictyostelium discoideum"
                     /mol_type="mRNA"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="dda23e04"
                     /sex="mat A"
                     /dev_stage="Aggregation stage"
                     /clone_lib="Dictyostelium discoideum cDNA library, AF"
ORIGIN

Query Match           3.8%;  Score 42.8;  DB 4;  Length 640;
Best Local Similarity 48.2%;  Pred. No. 0.68;
Matches  119;  Conservative   0;  Mismatches  128;  Indels   0;  Gaps   0;

Qy   412 AAAACAGTACAACGTCACTCAAGCATTGGGAGACGTGGTCCAGAACAAACCCAAGGAAA 471
                |   | ||     || ||||  ||| |||| || ||    | ||| ||  ||||
Db   275 AACAAAACAGGTGGTTAACAAATGATTTTGGATTACTTGGACCTGCCATTTACCACCAGTTAA 216

Qy   472 TTTCGGGACCAAGACCTAATCAGACAAGGAACTGATTCAGAACAGATTCAAACAATTGCCGCCAAATTGC 531
                |     ||| | ||  || ||   ||      | || ||  |||      |||||||  ||
Db   215 AGATTTGGAATCAAACCAAGCTTAAGTCAATGAATGTCAATGAATTGTAAGTCAATGAATGTTACGAAATTTAC 156

Qy   532 ACAATTGCTCCAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGATGAAGTCAC 591
                || ||||| ||||| ||   |  || ||||||| || | | | || ||
Db   155 CAAATTTGCACCAGTTGAGTGAAATCATTTGGAATTACAAACTCCAATTCTAACTGGGAA 96

Qy   592 ACCTTCGGGACCATGCGCTGACTTATCATGGAGCCATTAAATTGGATGACAAAGATCCACA 651
                || | | |    |  |  |  ||   |  | |   || ||||  || || |
Db    95 ACCTGGTTGTAAACCACTGAAGTCCATCTACCATTNAATCGGTTGTAACGGTTCCAAA 36

Qy   652 ATTCAAA 658
              |||
Db    35 GAGGAAA 29

RESULT 4
CNS00L00
LOCUS       CNS00L00                1101 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC:
```

Left column

ACCESSION AL068607
VERSION   AL068607.1 GI:4958689
KEYWORDS   GSS.
SOURCE   Drosophila melanogaster (fruit fly)
ORGANISM   Drosophila melanogaster
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
  Ephydroidea; Drosophilidae; Drosophila.

BACR32D23 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

REFERENCE 1 (bases 1 to 1101)
AUTHORS   Genoscope.
TITLE   Direct Submission
JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
  BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
  - Web : www.genoscope.cns.fr)

COMMENT   Determination of this BAC-end sequence was carried out as part of a
  collaboration with the Berkeley Drosophila Genome Project (BDGP).
  The BDGP is constructing a physical map of the Drosophila
  melanogaster genome using these BACs. For further information
  please see http://www.fruitfly.org The BDGP Drosophila
  melanogaster BAC library was prepared by Kazutoyo Osoegawa and
  Aaron Mammoser in Pieter de Jong's laboratory in the Department of
  Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
  NY. The library is named RPCI-98 and was constructed by partial
  EcoRI digestion of Drosophila DNA provided by the BDGP from the
  isogenic strain y2; cn bw sp, the same strain used for the BDGP's
  P1 and EST libraries. A more detailed description of the library
  and how to order individual BAC clones, the entire library, or
  filters for hybridization from the BACPAC Resource Center can be
  found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES   Location/Qualifiers
  source 1..1101
          /organism="Drosophila melanogaster"
          /mol_type="genomic DNA"
          /db_xref="taxon:7227"
          /clone="BACR32D23"
          /clone_lib="RPCI-98"
          /note="end : TET3"

ORIGIN
Query Match   3.7%; Score 42.4; DB 9; Length 1101;
Best Local Similarity 18.7%; Pred. No. 1.1;
Matches   75; Conservative 154; Mismatches 172; Indels 1; Gaps 1;

Qy   639 ACAAAGATCCACAATTCAAAGACAACGTCATACTGCTGAACAAGCACATTGACGCATACA 698
Db   690 MMMWAAMHTAMTCACHMMAAMMAMMTMMTTWWAMMACMRMAMMMCCCSCAMMCMAMAMAHW 749
Qy   699 AAACATTCCCACCACAGAGCCTAAAAAGGACAAAAAAGAGAAAAGAACTGATGAAGCTCAGC 758
Db   750 MACMAWHHMCMCMMMCMMMMCMMMMCAMAMWTACAMCAAMAAAMWWAYMMHCACMMMYCMMMM 809
Qy   759 CTTTGCCGCAGAGACAAAAGAGCAGCCCACTGTGACTCTTCTTCCTGCGGCTGACATGG 818
Db   810 AHCTMYACCMCMCMCYCACMYCWCCCMMMCMWAMTMMAMMAMTTTHCHTTTCMYHTTTTTY 869
Qy   819 ATGATTTCTCCAGACAACTTCAAAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAGG 878
Db   870 CMTMC--CCMMMMMMMMACHMAMMHCMMWTHMWHWAHAWTHMCMCCCCTMMWHAHTWHWHTHHH 928
Qy   879 CATAACACTCATGATGACCACACAAGGCAGCAGATGGGCTATGTAAACGTTTTCGCAATTCC 938
Db   929 HWMCAAMMMACHMAMMCMCCMMWMYCMWMCMMWTYTMHAMMHHTTTTTMMTHCMYMCA 988
Qy   939 GTTTACGATACATAGTCTACTCTTGTGCAGAATGAATTCTCGTAACTAAACAGCACAAGT 998
Db   989 MMGMMAWNAMCMMMAAMAMWTCMAAHAMDMAMTWTTTTTWHMAMAMAAWAAAAMMMMCY 1048
Qy   999 AGGTTTAGTTAACTTTAATCCACATAGCAATCTTTAATCAA 1040
Db 1049 WCAMTMAMMHHAYTMWTTHBMTNMHHAMAACMAMCMYMMMAM 1090

REFERENCE　　rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
　　　　　　　Glycine.
　　　　　　　1 (bases 1 to 518)
AUTHORS　　　Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
　　　　　　　Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
　　　　　　　Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
　　　　　　　Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
　　　　　　　Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
　　　　　　　McCann,R., Waterston,R. and Wilson,R.
TITLE　　　　Public Soybean EST Project
JOURNAL　　　Unpublished (1999)
COMMENT　　　Contact: Shoemaker R/Public Soybean EST Project
　　　　　　　Public Soybean EST Project
　　　　　　　Washington University School of Medicine
　　　　　　　4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
　　　　　　　Tel: 314 286 1800
　　　　　　　Fax: 314 286 1810
　　　　　　　Email: est@watson.wustl.edu
　　　　　　　When it has been determined, an EST from the other end of this
　　　　　　　clone is listed in the 'Other ESTs on clone' field. This clone is
　　　　　　　available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
　　　　　　　57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
　　　　　　　High quality sequence stop: 421.
FEATURES　　Location/Qualifiers
　　source　　1..518
　　　　　　　/organism="Glycine soja"
　　　　　　　/mol_type="mRNA"
　　　　　　　/db_xref="taxon:3848"
　　　　　　　/clone="GENOME SYSTEMS CLONE ID: Gm-c1057-187"
　　　　　　　/tissue_type="Degenerating cotyledons, 2 week old
　　　　　　　seedling"
　　　　　　　/lab_host="DH10B"
　　　　　　　/clone_lib="Gm-c1057"
　　　　　　　/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
　　　　　　　XhoI; The cDNA library was constructed from mRNA isolated
　　　　　　　from degenerating cotyledons of 2 week old seedlings from
　　　　　　　PI468916. Complementary DNA was synthesized from mRNA
　　　　　　　using a primer consisting of a poly(dT) sequence with a
　　　　　　　XhoI restriction site. EcoRI adapters were ligated to the
　　　　　　　blunt-ended cDNA fragments followed by XhoI digestion.
　　　　　　　The cDNA fragments were directionally cloned into the
　　　　　　　EcoRI-XhoI restriction site of the pBluescript vector. The
　　　　　　　ligated cDNA fragments were transformed into DH10B host
　　　　　　　cells (GibcoBRL). This library was constructed in the
　　　　　　　laboratory of Dr. Randy Shoemaker."
ORIGIN

Query Match　　　3.7%;　Score 42.2;　DB 2;　Length 518;
Best Local Similarity　54.1%;　Pred. No. 0.96;
Matches　86;　Conservative　0;　Mismatches　73;　Indels　0;　Gaps　0;

Qy 622 AGGCCATTAAATTGGATGAGACAAAGATCCACAATTCACAAGACAACGTCATACTGCTGAACAA 681
       |||| ||| |||| |||| |||| || ||||| |||
Db 332 AGCAAGCAATTAGTGACAACAAAGATTACAAAGATTACAGTAGTAAAACTTACCATCTACTGCAAAACGA 273

Qy 682 GCACATTGACGCATACAAAACATTCCCACCACAGAGCCTAAAAAGGACAAAAGAAAA 741
       || |||| ||| ||| |||||| |||||| ||||
Db 272 GCCAATCTAGAAGCTCCCATCCTTTTGACAACAGGGATCCTTGACAACGCAAAAGCGAAACA 213

Qy 742 GACTGATGAAGCTCAGCCTTTGCCGCGAGAGACAAAAGAA 780
       ||||| |||| | |||| ||||||
Db 212 GACTGTTGCATCAAATTCTCTCCCATATAAGACAAAGAAA 174

RESULT 7
BJ362480/c
LOCUS　　　　BJ362480　　540 bp　　mRNA　　linear　　EST 07-MAR-2002
DEFINITION　BJ362480 Dictyostelium discoideum cDNA library, CF Dictyostelium
　　　　　　　discoideum cDNA clone ddc22a02 5', mRNA sequence.
ACCESSION　　BJ362480
VERSION　　　BJ362480.1　GI:19262075
KEYWORDS　　EST.
SOURCE　　　Dictyostelium discoideum

ORGANISM　　Dictyostelium discoideum
　　　　　　　Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE　　1 (bases 1 to 540)
AUTHORS　　　Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
TITLE　　　　Full length cDNA of Dictyostelium discoideum at the culmination
　　　　　　　stage
JOURNAL　　　Unpublished (2002)
COMMENT　　　Contact: Tadasu Shin-i
　　　　　　　Center For Genetic Resource Information
　　　　　　　National Institute of Genetics
　　　　　　　1111 Yata, Mishima, Shizuoka 411-8540, Japan
　　　　　　　Tel: 81-559-81-6856
　　　　　　　Fax: 81-559-81-6855
　　　　　　　Email: tshini@genes.nig.ac.jp.
FEATURES　　Location/Qualifiers
　　source　　1..540
　　　　　　　/organism="Dictyostelium discoideum"
　　　　　　　/mol_type="mRNA"
　　　　　　　/strain="AX4"
　　　　　　　/db_xref="taxon:44689"
　　　　　　　/clone="ddc22a02"
　　　　　　　/sex="mat A"
　　　　　　　/dev_stage="Culmination stage"
　　　　　　　/clone_lib="Dictyostelium discoideum cDNA library, CF"

ORIGIN

Query Match　　　3.7%;　Score 42.2;　DB 4;　Length 540;
Best Local Similarity　48.2%;　Pred. No. 0.97;
Matches　119;　Conservative　0;　Mismatches　128;　Indels　0;　Gaps　0;

Qy 412 AAAACAGTACAACGTCACTCAAGCATTGGGAGACGTGGTCCAGAACAAACCCAAGGAAA 471
       ||| |||| |||| || ||||| ||| ||||
Db 527 AACAAAACAGGTGTGTAACAAATGATTTTGGATTACTTGGACCTGCATTTACACCAGTTAA 468

Qy 472 TTTCGGGACCAAGACCTAATCAGACAACAGGAACTGATTACAAACATTGGCCGCAAATTGC 531
       ||| |||| ||| ||| ||| |||| |||
Db 467 AGATTTGGAATCAAACCAATGCTTAAGTCAATAGATTGACCAACTGTTACGAAAATTTAC 408

Qy 532 ACAATTGCTCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCAAGTGGAAGTCAC 591
       |||| ||| ||| |||| |||||| |||||
Db 407 CAAATTTGCACCAGTTGAAGTGAAATCATTTGGAATTACAAACTCAATTCTAACTGGGAA 348

Qy 592 ACCTTCGGGACAACATGGCTGACTTATCATGGAGCCATTAAATTGGATGACAAAGATCCACA 651
       |||| ||||| ||||| ||||| |||| ||| ||||||
Db 347 ACCTGGTTGTGTAAACCACTGAATGAACCATCTATCATTGAATCGGTTGTAACGGTTCCAAA 288

Qy 652 ATTCAAA 658
       |||
Db 287 GAGGAAA 281

RESULT 8
BJ366773/c
LOCUS　　　　BJ366773　　556 bp　　mRNA　　linear　　EST 08-MAR-2002
DEFINITION　BJ366773 Dictyostelium discoideum cDNA library, CF Dictyostelium
　　　　　　　discoideum cDNA clone ddc40o04 5', mRNA sequence.
ACCESSION　　BJ366773
VERSION　　　BJ366773.1　GI:19276075
KEYWORDS　　EST.
SOURCE　　　Dictyostelium discoideum
ORGANISM　　Dictyostelium discoideum
　　　　　　　Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE　　1 (bases 1 to 556)
AUTHORS　　　Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
TITLE　　　　Full length cDNA of Dictyostelium discoideum at the culmination
　　　　　　　stage
JOURNAL　　　Unpublished (2002)
COMMENT　　　Contact: Tadasu Shin-i
　　　　　　　Center For Genetic Resource Information
　　　　　　　National Institute of Genetics
　　　　　　　1111 Yata, Mishima, Shizuoka 411-8540, Japan
　　　　　　　Tel: 81-559-81-6856
　　　　　　　Fax: 81-559-81-6855

```
FEATURES        Email: tshini@genes.nig.ac.jp.
  source        Location/Qualifiers
                1..556
                /organism="Dictyostelium discoideum"
                /mol_type="mRNA"
                /strain="AX4"
                /db_xref="taxon:44689"
                /clone="ddc40o04"
                /sex="mat A"
                /dev_stage="Culmination stage"
ORIGIN          /clone_lib="Dictyostelium discoideum cDNA library, CF"

Query Match           3.7%;   Score 42.2;   DB 4;   Length 556;
Best Local Similarity  48.2%;  Pred. No. 0.98;
Matches  119;  Conservative    0;  Mismatches  128;  Indels   0;   Gaps   0;

Qy  412 AAAACAGTACAACGTCACTCAAGCATTGGGAGACGTGGTCCAGAACAAACCCAAGGAAA 471
          ||||| ||  ||| ||  |||||| |||||| ||||  || |||||||||||||||||||
Db  457 AACAAAACAGTGGTAACAAATGATTTTGGATTACTTGGACCTGCATTACACCCAGTTAA 398

Qy  472 TTTCGGGGACCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAATTGC 531
          |    ||  | |||      ||||||       || ||  |||||||||||||||||||||
Db  397 AGATTTTGGAATCAAACCAATGCTTAAGTCAATAGATTGACCAACTGTTACGAAATTTAC 338

Qy  532 ACAATTTGCTCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCAC 591
          |||||||| ||  ||| ||   ||||||||||| ||||||
Db  337 CAAATTTGCACCAGTTGAAGTGAAATCATTTGGAATTACAAACTCAATTCAACTGGGAA 278

Qy  592 ACCTTCGGGAACATGGCTGACTTATCATGCAGCCCATTAAATTGGATGACAAAGATCCACA 651
          |||| ||  |  ||  ||  ||   |||| ||   ||||||  || ||||
Db  277 ACCTGGTTGTAAACCACTGAATGACCATCTACCATTTGAATCGGTTGTAACGGTTCCAAA 218

Qy  652 ATTCAAA 658
          |||
Db  217 GAGGAAA 211

RESULT 9
BJ328471/c
LOCUS           BJ328471          593 bp     mRNA       linear     EST 05-MAR-2002
DEFINITION      BJ328471 Dictyostelium discoideum cDNA library, AF Dictyostelium
                discoideum cDNA clone dda28e11 5', mRNA sequence.
ACCESSION       BJ328471
VERSION         BJ328471.1  GI:19158601
KEYWORDS        EST.
SOURCE          Dictyostelium discoideum
  ORGANISM      Dictyostelium discoideum
                Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
REFERENCE       1  (bases 1 to 593)
  AUTHORS       Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
  TITLE         Full length cDNA of Dictyostelium discoideum at the aggregation
                stage
  JOURNAL       Unpublished (2002)
COMMENT         Contact: Tadasu Shin-i
                Center For Genetic Resource Information
                National Institute of Genetics
                1111 Yata, Mishima, Shizuoka 411-8540, Japan
                Tel: 81-559-81-6856
                Fax: 81-559-81-6855
                Email: tshini@genes.nig.ac.jp.
FEATURES        Location/Qualifiers
  source        1..593
                /organism="Dictyostelium discoideum"
                /mol_type="mRNA"
                /strain="AX4"
                /db_xref="taxon:44689"
                /clone="dda28e11"
                /sex="mat A"
                /dev_stage="Aggregation stage"
ORIGIN          /clone_lib="Dictyostelium discoideum cDNA library, AF"
```

```
Query Match           3.7%;   Score 42.2;   DB 4;   Length 593;
Best Local Similarity  48.2%;  Pred. No. 1;
Matches  119;  Conservative    0;  Mismatches  128;  Indels   0;   Gaps   0;

Qy  412 AAAACAGTACAACGTCACTCAAGCATTGGGAGACGTGGTCCAGAACAAACCCAAGGAAA 471
          ||||| ||  ||| ||  |||||| |||||| ||||  || |||||||||||||||||||
Db  303 AACAAAACAGTGGTAACAAATGATTTTGGATTACTTGGACCTGCATTACACCCAGTTAA 244

Qy  472 TTTCGGGGACCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAATTGC 531
          |    ||  | |||      ||||||       || ||  |||||||||||||||||||||
Db  243 AGATTTTGGAATCAAACCAATGCTTAAGTCAATAGATTGACCAACTGTTACGAAATTTAC 184

Qy  532 ACAATTTGCTCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCAC 591
          |||||||| ||  ||| ||   ||||||||||| ||||||
Db  183 CAAATTTGCACCAGTTGAAGTGAAATCATTTGGAATTACAAACTCAATTCAACTGGGAA 124

Qy  592 ACCTTCGGGAACATGGCTGACTTATCATGCAGCCCATTAAATTGGATGACAAAGATCCACA 651
          |||| ||  |  ||  ||  ||   |||| ||   ||||||  || ||||
Db  123 ACCTGGTTGTAAACCACTGAATGACCATCTACCATTTGAATCGGTTGTAACGGTTCCAAA 64

Qy  652 ATTCAAA 658
          |||
Db   63 GAGGAAA 57

RESULT 10
BJ339144/c
LOCUS           BJ339144          601 bp     mRNA       linear     EST 07-MAR-2002
DEFINITION      BJ339144 Dictyostelium discoideum cDNA library, AF Dictyostelium
                discoideum cDNA clone dda64f10 5', mRNA sequence.
ACCESSION       BJ339144
VERSION         BJ339144.1  GI:19247506
KEYWORDS        EST.
SOURCE          Dictyostelium discoideum
  ORGANISM      Dictyostelium discoideum
                Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
REFERENCE       1  (bases 1 to 601)
  AUTHORS       Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
  TITLE         Full length cDNA of Dictyostelium discoideum at the aggregation
                stage
  JOURNAL       Unpublished (2002)
COMMENT         Contact: Tadasu Shin-i
                Center For Genetic Resource Information
                National Institute of Genetics
                1111 Yata, Mishima, Shizuoka 411-8540, Japan
                Tel: 81-559-81-6856
                Fax: 81-559-81-6855
                Email: tshini@genes.nig.ac.jp.
FEATURES        Location/Qualifiers
  source        1..601
                /organism="Dictyostelium discoideum"
                /mol_type="mRNA"
                /strain="AX4"
                /db_xref="taxon:44689"
                /clone="dda64f10"
                /sex="mat A"
                /dev_stage="Aggregation stage"
                /clone_lib="Dictyostelium discoideum cDNA library, AF"
ORIGIN

Query Match           3.7%;   Score 42.2;   DB 4;   Length 601;
Best Local Similarity  48.2%;  Pred. No. 1;
Matches  119;  Conservative    0;  Mismatches  128;  Indels   0;   Gaps   0;

Qy  412 AAAACAGTACAACGTCACTCAAGCATTGGGAGACGTGGTCCAGAACAAACCCAAGGAAA 471
          ||||| ||  ||| ||  |||||| |||||| ||||  || |||||||||||||||||||
Db  582 AACAAAACAGTGGTAACAAATGATTTTGGATTACTTGGACCTGCATTACACCCAGTTAA 523

Qy  472 TTTCGGGGACCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAATTGC 531
          |    ||  | |||      ||||||       || ||  |||||||||||||||||||||
Db  522 AGATTTTGGAATCAAACCAATGCTTAAGTCAATAGATTGACCAACTGTTACGAAATTTAC 463

Qy  532 ACAATTTGCTCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCAC 591
```

Db   462 CAAATTTGCACCAGTTGAAGTGAAATCATTTGGAATTACAAACTCAATTCTAACTGGGAA 403
          |||||||| ||| || |||||| ||
Qy   592 ACCTTCGGGAACATGGCTGACTTATCATGGAGCCATTAAATTGGATGACAAAGATCCACA 651
          |||| || || |||| ||| ||||||||
Db   402 ACCTGGTTGTAAACCACTGAATGACCATCTACCATTTGAATCGGTTGTAACGGTTCCAAA 343

Qy   652 ATTCAAA 658
          |||
Db   342 GAGGAAA 336

RESULT 11
BJ365525/c
LOCUS   BJ365525   605 bp   mRNA   linear   EST 08-MAR-2002
DEFINITION BJ365525 Dictyostelium discoideum cDNA library, CF Dictyostelium
            discoideum cDNA clone ddc35m13 5', mRNA sequence.
ACCESSION   BJ365525
VERSION   BJ365525.1 GI:19274827
KEYWORDS   EST.
SOURCE   Dictyostelium discoideum
  ORGANISM Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
REFERENCE   1 (bases 1 to 605)
  AUTHORS   Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
  TITLE   Full length cDNA of Dictyostelium discoideum at the culmination
            stage
  JOURNAL   Unpublished (2002)
COMMENT   Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES   Location/Qualifiers
     source   1..605
                     /organism="Dictyostelium discoideum"
                     /mol_type="mRNA"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="ddc35m13"
                     /sex="mat A"
                     /dev_stage="Culmination stage"
                     /clone_lib="Dictyostelium discoideum cDNA library, CF"
ORIGIN

Query Match   3.7%;   Score 42.2;   DB 4;   Length 605;
Best Local Similarity   48.2%;   Pred. No. 1;
Matches 119;   Conservative   0;   Mismatches   128;   Indels   0;   Gaps   0;

Qy   412 AAAACAGTACAACGTCACTCAAGCATTGGGAGACGTGGTCCAGAACAAACCCAAGGAAA 471
          ||||||||||||||||||| || || |||||||||||||
Db   509 AACAAAACAGGTGGTAACAAATGATTTTGGATTACTTGGACCTGCATTTACACCAGTTAA 450

Qy   472 TTTCGGGACCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAATTGC 531
          ||||| ||||| ||| ||
Db   449 AGATTTTGGAATCAAACCAATGCTTAAGTCAATAGATTGACCAACTGTTACGAAATTTAC 390

Qy   532 ACAATTTGCTCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCAC 591
          ||||||||| ||| || |||| || | || || ||||||
Db   389 CAAATTTGCACCAGTTGAAGTGAAATCATTTGGAATTACAAACTCAATTCTAACTGGGAA 330

Qy   592 ACCTTCGGGAACATGGCTGACTTATCATGGAGCCATTAAATTGGATGACAAAGATCCACA 651
          |||| || || |||| ||| ||||||||
Db   329 ACCTGGTTGTAAACCACTGAATGACCATCTACCATTTGAATCGGTTGTAACGGTTCCAAA 270

Qy   652 ATTCAAA 658
          |||
Db   269 GAGGAAA 263

RESULT 12

Db   462 CAAATTTGCACCAGTTGAAGTGAAATCATTTGGAATTACAAACTCAATTCTAACTGGGAA 403
          ||||||| || || |||||| |||||||
Qy   592 ACCTTCGGGAACATGGCTGACTTATCATGGAGCCATTAAATTGGATGACAAAGATCCACA 651
          |||| || || |||| ||| ||||||||
Db   402 ACCTGGTTGTAAACCACTGAATGACCATCTACCATTTGAATCGGTTGTAACGGTTCCAAA 343

Qy   652 ATTCAAA 658
          |||
Db   342 GAGGAAA 336

RESULT 11
BJ339537/c
LOCUS   BJ339537   606 bp   mRNA   linear   EST 07-MAR-2002
DEFINITION BJ339537 Dictyostelium discoideum cDNA library, AF Dictyostelium
            discoideum cDNA clone dda65o17 5', mRNA sequence.
ACCESSION   BJ339537
VERSION   BJ339537.1 GI:19247899
KEYWORDS   EST.
SOURCE   Dictyostelium discoideum
  ORGANISM Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
REFERENCE   1 (bases 1 to 606)
  AUTHORS   Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
  TITLE   Full length cDNA of Dictyostelium discoideum at the aggregation
            stage
  JOURNAL   Unpublished (2002)
COMMENT   Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES   Location/Qualifiers
     source   1..606
                     /organism="Dictyostelium discoideum"
                     /mol_type="mRNA"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="dda65o17"
                     /sex="mat A"
                     /dev_stage="Aggregation stage"
                     /clone_lib="Dictyostelium discoideum cDNA library, AF"
ORIGIN

Query Match   3.7%;   Score 42.2;   DB 4;   Length 606;
Best Local Similarity   48.2%;   Pred. No. 1;
Matches 119;   Conservative   0;   Mismatches   128;   Indels   0;   Gaps   0;

Qy   412 AAAACAGTACAACGTCACTCAAGCATTGGGAGACGTGGTCCAGAACAAACCCAAGGAAA 471
          ||||||||||||||||||| || || |||||||||||||
Db   508 AACAAAACAGGTGGTAACAAATGATTTTGGATTACTTGGACCTGCATTTACACCAGTTAA 449

Qy   472 TTTCGGGACCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAATTGC 531
          ||||| ||||| ||| ||
Db   448 AGATTTTGGAATCAAACCAATGCTTAAGTCAATAGATTGACCAACTGTTACGAAATTTAC 389

Qy   532 ACAATTTGCTCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCAC 591
          ||||||||| ||| || |||| || | || || ||||||
Db   388 CAAATTTGCACCAGTTGAAGTGAAATCATTTGGAATTACAAACTCAATTCTAACTGGGAA 329

Qy   592 ACCTTCGGGAACATGGCTGACTTATCATGGAGCCATTAAATTGGATGACAAAGATCCACA 651
          |||| || || |||| ||| ||||||||
Db   328 ACCTGGTTGTAAACCACTGAATGACCATCTACCATTTGAATCGGTTGTAACGGTTCCAAA 269

Qy   652 ATTCAAA 658
          |||
Db   268 GAGGAAA 262

RESULT 13
BJ387757/c
LOCUS   BJ387757   616 bp   mRNA   linear   EST 08-MAR-2002
DEFINITION BJ387757 Dictyostelium discoideum cDNA library, SF Dictyostelium
            discoideum cDNA clone dds4a23 5', mRNA sequence.
ACCESSION   BJ387757
VERSION   BJ387757.1 GI:19297141
KEYWORDS   EST.
SOURCE   Dictyostelium discoideum
  ORGANISM Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
REFERENCE   1 (bases 1 to 616)
  AUTHORS   Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
  TITLE   Full length cDNA of Dictyostelium discoideum at the slug stage

```
JOURNAL     Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp
FEATURES
     source     1..616
                /organism="Dictyostelium discoideum"
                /mol_type="mRNA"
                /strain="AX4"
                /db_xref="taxon:44689"
                /clone="dds4a23"
                /sex="mat A"
                /dev_stage="Slug stage"
                /clone_lib="Dictyostelium discoideum cDNA library, SF"
ORIGIN

Query Match             3.7%;  Score 42.2;  DB 4;  Length 616;
Best Local Similarity  48.2%;  Pred. No. 1;
Matches  119;  Conservative   0;  Mismatches 128;  Indels  0;  Gaps  0;

Qy    412 AAAACAGTACAACGTCACTCAAGCATTGGGAGACGTGGTCCAGAACAACCCAAGGAAA 471
Db    516 AACAAAACAGGTGGTAACAAATGATTTTGGATTACTTGGACCTGCATTTACACCAGTTAA 457
Qy    472 TTTCGGGGACCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAATTGC 531
Db    456 AGATTTTGGAATCAAACCAATGCTTAAGTCAATAGATTGACCAACTGTTACGAAATTTAC 397
Qy    532 ACAATTTGCTCCAAGTGCCTCTGCCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCAC 591
Db    396 CAAATTTGCACCAGTTGAAGTGAAATCATTTGGAATTACAAACTCAATTCTAACTGGGAA 337
Qy    592 ACCTTCGGGAACATGCTGACTTATCATGGACCATTAAATTGGATGACAAAGATCCACA 651
Db    336 ACCTGGTTGTAAACCACTGAATGACCATCTACCATTTGAATCGGTTGTAACGGTTCCAAA 277
Qy    652 ATTCAAA 658
Db    276 GAGGAAA 270

RESULT 14
BJ366277/c
LOCUS       BJ366277                617 bp    mRNA    linear    EST 08-MAR-2002
DEFINITION  BJ366277 Dictyostelium discoideum cDNA library, CF Dictyostelium
            discoideum clone ddc38112 5', mRNA sequence.
ACCESSION   BJ366277
VERSION     BJ366277.1  GI:19275579
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
REFERENCE   1  (bases 1 to 617)
  AUTHORS   Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
  TITLE     Full length cDNA of Dictyostelium discoideum at the culmination
            stage
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp
FEATURES
     source     1..617
                /organism="Dictyostelium discoideum"
                /mol_type="mRNA"
                /strain="AX4"
                /db_xref="taxon:44689"
                /clone="ddc38112"
                /sex="mat A"
                /dev_stage="Culmination stage"
                /clone_lib="Dictyostelium discoideum cDNA library, CF"
ORIGIN

Query Match             3.7%;  Score 42.2;  DB 4;  Length 617;
Best Local Similarity  48.2%;  Pred. No. 1;
Matches  119;  Conservative   0;  Mismatches 128;  Indels  0;  Gaps  0;

Qy    412 AAAACAGTACAACGTCACTCAAGCATTGGGAGACGTGGTCCAGAACAACCCAAGGAAA 471
Db    518 AACAAAACAGGTGGTAACAAATGATTTTGGATTACTTGGACCTGCATTTACACCAGTTAA 459
Qy    472 TTTCGGGGACCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAATTGC 531
Db    458 AGATTTTGGAATCAAACCAATGCTTAAGTCAATAGATTGACCAACTGTTACGAAATTTAC 399
Qy    532 ACAATTTGCTCCAAGTGCCTCTGCCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCAC 591
Db    398 CAAATTTGCACCAGTTGAAGTGAAATCATTTGGAATTACAAACTCAATTCTAACTGGGAA 339
Qy    592 ACCTTCGGGAACATGCTGACTTATCATGGACCATTAAATTGGATGACAAAGATCCACA 651
Db    338 ACCTGGTTGTAAACCACTGAATGACCATCTACCATTTGAATCGGTTGTAACGGTTCCAAA 279
Qy    652 ATTCAAA 658
Db    278 GAGGAAA 272

RESULT 15
BJ370135/c
LOCUS       BJ370135                629 bp    mRNA    linear    EST 08-MAR-2002
DEFINITION  BJ370135 Dictyostelium discoideum cDNA library, CF Dictyostelium
            discoideum clone ddc53d06 5', mRNA sequence.
ACCESSION   BJ370135
VERSION     BJ370135.1  GI:19279518
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
REFERENCE   1  (bases 1 to 629)
  AUTHORS   Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
  TITLE     Full length cDNA of Dictyostelium discoideum at the culmination
            stage
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp
FEATURES
     source     1..629
                /organism="Dictyostelium discoideum"
                /mol_type="mRNA"
                /strain="AX4"
                /db_xref="taxon:44689"
                /clone="ddc53d06"
                /sex="mat A"
                /dev_stage="Culmination stage"
                /clone_lib="Dictyostelium discoideum cDNA library, CF"
ORIGIN

Query Match             3.7%;  Score 42.2;  DB 4;  Length 629;
Best Local Similarity  48.2%;  Pred. No. 1;
Matches  119;  Conservative   0;  Mismatches 128;  Indels  0;  Gaps  0;

Qy    412 AAAACAGTACAACGTCACTCAAGCATTGGGAGACGTGGTCCAGAACAACCCAAGGAAA 471
```

```
Db   571 AACAAAACAGGTCGTAACAGATGATTTTGGATTACTTCGACCTGCATTTACACCAGTTAA 512

Qy   472 TTTCGGGGACCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAATTGC 531
             ||  |  |||| || | |||| | | |||| ||||||  |  | | |||  |  |  ||
Db   511 AGATTTTGGAATCAAACCAATGCTTAAGTCAATAGATTGACCAACTGTTACGAAATTTAC 452

Qy   532 ACAATTTGCTCCAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCAC 591
             ||||||| ||| || ||   || |  ||||||||| ||   ||      |  | | |
Db   451 CAAATTTGCACCAGTTGAAGTGAAATCATTTGGAATTACAAACTCAATTCTAACTGGGAA 392

Qy   592 ACCTTCGGGAACATGGCTGACTTATCATGGAGCCATTAAATTCGATGACAAAGATCCACA 651
             |||| |  |   |||  | |||  ||| || | ||||  | |||| | || || |||| |
Db   391 ACCTGGTTGTAAACCACTGAATGACCATCTACCATTTGAATCGGTTGTAACGGTTCCAAA 332

Qy   652 ATTCAAA 658
             | ||
Db   331 GAGGAAA 325
```

Search completed: May 15, 2005, 14:18:37
Job time : 5008.54 secs

```
              GenCore version 5.1.6
        Copyright (c) 1993 - 2005  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:      May 15, 2005, 11:00:45 ; Search time 95.9067 Seconds
                                    (without alignments)
                                 10609.895 Million cell updates/sec

Title:        US-10-764-075-2
Perfect score: 21
Sequence:      1 atgaccacacaaggcagatgg 21

Scoring table: IDENTITY NUC
               Gapop 10.0 , Gapext 1.0

Searched:     4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters:     9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    GenEmbl:*
              1: gb_ba:*
              2: gb_htg:*
              3: gb_in:*
              4: gb_om:*
              5: gb_ov:*
              6: gb_pat:*
              7: gb_ph:*
              8: gb_pl:*
              9: gb_pr:*
             10: gb_ro:*
             11: gb_sts:*
             12: gb_sy:*
             13: gb_un:*
             14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                         SUMMARIES

Result       Query
No.   Score  Match Length DB  ID        Description
----- -----  ----- ------ --  --        -----------
  1    21    100.0  1669  14  AY536760   AY536760 SARS coro
  2    21    100.0  1873  14  AY534758S4 AY534761 SARS coro
  3    21    100.0  1938  14  AY534762S6 AY534767 SARS coro
  4    21    100.0  2304  14  AY322205S4 AY322208 SARS coro
  5    21    100.0  2810  14  AY290752   AY290752 SARS coro
  6    21    100.0  8439  14  AY304489   AY304489 SARS coro
  7    21    100.0  8581  14  AY304487   AY304487 SARS coro
  8    21    100.0 11006  14  AY304491   AY304491 SARS coro
  9    21    100.0 11010  14  AY304493   AY304493 SARS coro
 10    21    100.0 11010  14  AY304494   AY304494 SARS coro
 11    21    100.0 13471  14  AY304490   AY304490 SARS coro
 12    21    100.0 13471  14  AY304492   AY304492 SARS coro
 13    21    100.0 29013  14  AY463060   AY463060 SARS coro
 14    21    100.0 29350  14  AY394999   AY394999 SARS coro
 15    21    100.0 29350  14  AY395000   AY395000 SARS coro
 16    21    100.0 29350  14  AY395001   AY395001 SARS coro
 17    21    100.0 29350  14  AY395002   AY395002 SARS coro
 18    21    100.0 29433  14  AY394977   AY394977 SARS coro
 19    21    100.0 29530  14  AY394985   AY394985 SARS coro
 20    21    100.0 29573  14  AY338174   AY338174 SARS coro
 21    21    100.0 29573  14  AY338175   AY338175 SARS coro
 22    21    100.0 29573  14  AY348314   AY348314 SARS coro
 23    21    100.0 29577  14  AY559094   AY559094 SARS coro
 24    21    100.0 29592  14  AY463059   AY463059 SARS coro
 25    21    100.0 29620  14  AY395004   AY395004 SARS coro
 26    21    100.0 29640  14  AY394978   AY394978 SARS coro
 27    21    100.0 29645  14  AY394979   AY394979 SARS coro
 28    21    100.0 29646  14  AY394982   AY394982 SARS coro
 29    21    100.0 29647  14  AY395003   AY395003 SARS coro
 30    21    100.0 29661  14  AY559086   AY559086 SARS coro
 31    21    100.0 29665  14  AY394988   AY394988 SARS coro
 32    21    100.0 29670  14  AY559082   AY559082 SARS coro
 33    21    100.0 29683  14  AY394996   AY394996 SARS coro
 34    21    100.0 29683  14  AY394997   AY394997 SARS coro
 35    21    100.0 29699  14  AY394983   AY394983 SARS coro
 36    21    100.0 29705  14  AY283795   AY283795 SARS coro
 37    21    100.0 29705  14  AY394980   AY394980 SARS coro
 38    21    100.0 29706  14  AY283797   AY283797 SARS coro
 39    21    100.0 29709  14  AY394987   AY394987 SARS coro
 40    21    100.0 29710  14  AY559091   AY559091 SARS coro
 41    21    100.0 29711  14  AY283794   AY283794 SARS coro
 42    21    100.0 29711  14  AY283796   AY283796 SARS coro
 43    21    100.0 29711  14  AY283798   AY283798 SARS coro
 44    21    100.0 29711  14  AY427439   AY427439 SARS coro
 45    21    100.0 29712  14  AY559093   AY559093 SARS coro

                         ALIGNMENTS

RESULT 1
AY536760
LOCUS       AY536760               1669 bp    mRNA    linear   VRL 15-SEP-2004
DEFINITION  SARS coronavirus BJ01 nucleocapsid protein mRNA, complete cds.
ACCESSION   AY536760
VERSION     AY536760.3  GI:52100973
KEYWORDS    
SOURCE      SARS coronavirus BJ01
  ORGANISM  SARS coronavirus BJ01
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1 (bases 1 to 1669)
  AUTHORS   Li,T., Li,X., Chang,Z. and Liu,L.
  TITLE     Identification of SARS-CoV mRNA leader sequence
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 1669)
  AUTHORS   Li,T., Li,X., Liu,L. and Chang,Z.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JAN-2004) Institute of Biomedicine,Tsinghua
            University, Beijing 100084, China
REFERENCE   3 (bases 1 to 1669)
  AUTHORS   Li,T., Li,X., Liu,L. and Chang,Z.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUL-2004) Institute of Biomedicine,Tsinghua
            University, Beijing 100084, China
  REMARK    Sequence update by submitter
REFERENCE   4 (bases 1 to 1669)
  AUTHORS   Li,T., Li,X., Chang,Z. and Liu,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-SEP-2004) Institute of Biomedicine, Tsinghua
            University, Beijing 100084, China
  REMARK    Sequence update by submitter
  COMMENT   On Sep 15, 2004 this sequence version replaced gi:49921010.
FEATURES             Location/Qualifiers
     source          1..1669
                     /organism="SARS coronavirus BJ01"
                     /mol_type="mRNA"
                     /strain="BJ01"
                     /db_xref="taxon:228407"
                     /country="China"
     CDS             81..1349
                     /codon_start=1
```

```
                    /product="nucleocapsid protein"
                    /protein_id="AAS48456.1"
                    /db_xref="GI:44889746"
                    /translation="MSDNGPQSNQRSAPRITFGGPTDSTDNNQNGRNGARPKQRRPQ
          GLPNNTASWFTALTQHGKEELRFPRGQGVPINTNSGPDDQIGYYRRATRRVRGDGKM
          KELSPRWYFYLGTGPEASLPYGANKEGIVWVATEGALNTPKDHIGTRNPNNNAATVL
          QLPQGTTLPKGFYAEGSRGGSQASSRSSRGNSRNSTPGSSRGNSPARMASGGGET
          ALALLLDRLNQLESKVSGKGQQQQGQTVTKKSAAEASKKPRQKRTATKQYNVTQAFG
          RRGPEQTQGNFGDQDLLRQGTDYKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTY
          HGAIKLDDKDPQFKDNVILLNKHIDAYKTFPPTEPKDKKKKTDEAQPLPQRQKKQPT
          VTPLPAADMDDFSRQLQNSMSGASADSTQA"
ORIGIN

Query Match       100.0%;  Score 21;  DB 14;  Length 1669;
Best Local Similarity  100.0%;  Pred. No. 0.62;
Matches  21;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy         1 ATGACCACCACAAGGCAGATGG 21
             |||||||||||||||||||||
Db      1359 ATGACCACCACAAGGCAGATGG 1379


RESULT 2
AY534758S4
LOCUS       AY534758S4       1873 bp    RNA    linear    VRL 17-MAR-2004
DEFINITION  SARS coronavirus Sin0409, partial sequence.
ACCESSION   AY534761
VERSION     AY534761.1  GI:45384968
KEYWORDS    .
SEGMENT     4 of 4
SOURCE      SARS coronavirus Sin0409
  ORGANISM  SARS coronavirus Sin0409
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 1873)
  AUTHORS   Lim,P.L., Kurup,A., Gopalakrishna,G., Chan,K.P., Wong,C.W.,
            Ng,L.C., Se-Thoe,S.Y., Oon,L., Bai,X., Stanton,L.W., Ruan,Y.,
            Miller,L.D., Vega,V.B., James,L., Ooi,P.L., Kai,C.S., Olsen,S.J.,
            Ang,B. and Leo,Y.S.
  TITLE     Laboratory-acquired severe acute respiratory syndrome (SARS) -
            Singapore 2003
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1873)
  AUTHORS   Wei,C.L., Lee,C., Lin,S., Thoreau,H., Vega,V.B., Stanton,L.W. and
            Ruan,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JAN-2004) Genome Institute of Singapore, 60, Biopolis
            Street, 02-01, Genome, Singapore 138672, Singapore
FEATURES             Location/Qualifiers
     source          1..1873
                     /organism="SARS coronavirus Sin0409"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:266147"
                     /country="Singapore"
ORIGIN

Query Match       100.0%;  Score 21;  DB 14;  Length 1873;
Best Local Similarity  100.0%;  Pred. No. 0.62;
Matches  21;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy         1 ATGACCACCACAAGGCAGATGG 21
             |||||||||||||||||||||
Db      1598 ATGACCACCACAAGGCAGATGG 1618


RESULT 3
AY534762S6
LOCUS       AY534762S6       1938 bp    RNA    linear    VRL 17-MAR-2004
DEFINITION  SARS coronavirus Sin_WNV, partial sequence.
ACCESSION   AY534767
VERSION     AY534767.1  GI:45384975
KEYWORDS    .
SEGMENT     6 of 6
```

```
SOURCE      SARS coronavirus Sin_WNV
  ORGANISM  SARS coronavirus Sin_WNV
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 1938)
  AUTHORS   Lim,P.L., Kurup,A., Gopalakrishna,G., Chan,K.P., Wong,C.W.,
            Ng,L.C., Se-Thoe,S.Y., Oon,L., Bai,X., Stanton,L.W., Ruan,Y.,
            Miller,L.D., Vega,V.B., James,L., Ooi,P.L., Kai,C.S., Olsen,S.J.,
            Ang,B. and Leo,Y.S.
  TITLE     Laboratory-acquired severe acute respiratory syndrome (SARS) -
            Singapore 2003
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1938)
  AUTHORS   Wei,C.L., Lee,C., Lin,S., Thoreau,H., Vega,V.B., Stanton,L.W. and
            Ruan,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JAN-2004) Genome Institute of Singapore, 60, Biopolis
            Street, 02-01, Genome, Singapore 138672, Singapore
FEATURES             Location/Qualifiers
     source          1..1938
                     /organism="SARS coronavirus Sin_WNV"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:266148"
                     /country="Singapore"
ORIGIN

Query Match       100.0%;  Score 21;  DB 14;  Length 1938;
Best Local Similarity  100.0%;  Pred. No. 0.62;
Matches  21;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy         1 ATGACCACCACAAGGCAGATGG 21
             |||||||||||||||||||||
Db      1624 ATGACCACCACAAGGCAGATGG 1644


RESULT 4
AY322205S4
LOCUS       AY322205S4       2304 bp    RNA    linear    VRL 21-JUL-2003
DEFINITION  SARS coronavirus Shanhgai LY Orf7a gene, partial cds; and Orf7b,
            Orf8A, Orf8b, and nucleocapsid protein genes, complete cds.
ACCESSION   AY322208
VERSION     AY322208.1  GI:32454342
KEYWORDS    .
SEGMENT     4 of 4
SOURCE      SARS coronavirus Shanhgai LY
  ORGANISM  SARS coronavirus Shanhgai LY
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 2304)
  AUTHORS   Yuan,Z., Zhang,X., Hu,Y., Lan,S., Wang,H., Zhou,Z. and Wen,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUN-2003) Molecular Virology, Shanghai Medical
            College of Fudan University, 138 Yi Xue Yuan Road, Shanghai 200032,
            P.R. China
FEATURES             Location/Qualifiers
     source          1..2304
                     /organism="SARS coronavirus Shanhgai LY"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:235173"
                     /country="China: Shanghai"
     CDS             <1..276
                     /codon_start=1
                     /product="Orf7a"
                     /protein_id="AAP82970.1"
                     /db_xref="GI:32454348"
                     /translation="KEPCPSAIYEGNSPFHPLADNKFALTCTSTHFAFACADGTRHTY
                     QLRARSVSPKLFIRQEEVQQELYSPLFLIVAALVFLILCFTIKRKTE"
     CDS             273..407
                     /codon_start=1
                     /product="Orf7b"
                     /protein_id="AAP82971.1"
                     /db_xref="GI:32454349"
                     /translation="MNELTLIDFYLCFLAFLLFLVLIMLIIFWFSLEIQDLEEPCTKV
```

```
REFERENCE                 of Disease Prevention and Control, 17 Laozheda Rd., Hangzhou,
                          Zhejiang 310009, China
  AUTHORS       4 (bases 1 to 2810)
  TITLE         Wang,Z., Cheng,S., Zhang,Y., Lu,Y. and Mei.L.
  JOURNAL       Direct Submission
                Submitted (05-DEC-2003) Department of Microbiology, Zhejiang Center
                of Disease Prevention and Control, 17 Laozheda Rd., Hangzhou,
                Zhejiang 310009, China
  REMARK        Nucleotide and amino acid sequences updated by submitter
  COMMENT       On Dec 5, 2003 this sequence version replaced gi:3150596969.
  FEATURES             Location/Qualifiers
     source      1..2810
                 /organism="SARS coronavirus ZJ01"
                 /mol_type="genomic RNA"
                 /isolate="ZJ01b"
                 /isolation_source="patient with severe acute respiratory
                 syndrome"
                 /db_xref="taxon:230471"
                 /country="China: Hangzhou"
     misc_feature    1..6
                 /note="transcription regulatory core sequence 6"
     CDS         162..353
                 /codon_start=1
                 /product="uncharacterized protein 6"
                 /protein_id="AAR27473.1"
                 /db_xref="GI:38735510"
                 /translation="MFHLVDFQVTIAEILIIIMRTFRIAIWNLDVIISSIVRQLFKPL
                 TKKNYSELDDEEPMELDYP"
     misc_feature    331
                 /note="Xiaohong genotype locus 6"
     misc_feature    355..360
                 /note="transcription regulatory core sequence 7"
     CDS         361..729
                 /codon_start=1
                 /product="uncharacterized protein 7a"
                 /protein_id="AAR27474.1"
                 /db_xref="GI:38735511"
                 /translation="MKIILFLTLIVFTSCELYHYQECVRGTTVLLKEPCPSGTYEGNS
                 PFHPLADNKPALTCSTHFAFACADGTRHTYQLRARSVSPKLFIRQEEVQQELYSPLF
                 LIVAALVFLILCFTIKRKTE"
     CDS         726..860
                 /codon_start=1
                 /product="uncharacterized protein 7b"
                 /protein_id="AAR27475.1"
                 /db_xref="GI:38735512"
                 /translation="MNELTLIDFYLCFLPLAFLLFLVLIMLIIFWFSLEIQDLEEPCTKV"
     misc_feature    861..866
                 /note="transcription regulatory core sequence 8"
     CDS         867..986
                 /codon_start=1
                 /product="uncharacterized protein 8a"
                 /protein_id="AAR27476.1"
                 /db_xref="GI:38735513"
                 /translation="MCLKILVLRYNTRGNTYSTAWLCALGKVLPFHRWHTMVQTCTPNV
                 TINCQDPAGGALIARCWYLHEGHQTAAFRDVLVVLNKRTN"
     CDS         915
                 /translation="MKLLIVLTCISLCSCICTVVQRCASNKPHVLEDPCKVQH"
     misc_feature    952..1206
                 /note="Xiaohong genotype locus 7"
     CDS         952..1206
                 /codon_start=1
                 /product="uncharacterized protein 8b"
                 /protein_id="AAR27477.1"
                 /db_xref="GI:38735514"
     misc_feature    1193..1198
                 /note="transcription regulatory core sequence 9"
     CDS         1208..2476
                 /codon_start=1
                 /product="nucleocapsid protein"
                 /protein_id="AAP44772.1"
                 /db_xref="GI:31505970"
                 /translation="MSDNGPQSNQRSAPRITFGGPTDSTDNQNGGRNGARPKQRRPQ
                 GLPNNTASWFTALTQHGKEELRFPRGQGVPINTNSGPDDQIGYRRATRRVRGGDGKM
```

```
CDS
              "
              414..533
              /codon_start=1
              /product="Orf8A"
              /protein_id="AAP82972.1"
              /db_xref="GI:32454350"
              /translation="MKLLIVLTCISLCSCICTVVQRCASNKPHVLEDPCKVQH"
CDS           499..753
              /codon_start=1
              /product="Orf8b"
              /protein_id="AAP82973.1"
              /db_xref="GI:32454351"
              /translation="MCLKILVRYNTRGNTYSTAWLCALGKVLsPFHRWHTMVQTCTPNV
              TINCQDPAGGALIARCWYLHEGHQTAAFRDVLVVLNKRTN"
CDS           755..2023
              /codon_start=1
              /product="nucleocapsid protein"
              /protein_id="AAP82974.1"
              /db_xref="GI:32454352"
              /translation="MSDNGPQSNQRSAPRITFGGPTDSTDNQNGGRNGARPKQRRPQ
              GLPNNTASWFTALTQHGKEELRFPRGQGVPINTNSGPDDQIGYRRATRRVRGGDGKM
              KELSPRWYFYYLGTGPEASLPYGANKEGIVWVATEGALNTPKDHIGTRNPNNNAATVL
              QLPQGTTLPKGFYAEGSRGGSQASSRSSSRSRNSTPGSSRGNSPARMASGGGET
              ALALLLLDRLNQLESKVSGKGQQQQGQTVTKKSAAEASKKPRQKRTATKQYNVTQAFG
              RRGPEQTQGNFGDQDLIRQGTDYKHWPOIAQFPAPSASAFFGMSRIGMEAAPSGTWLTY
              HGAIKLDDKDPQFKDNVILLNKHIDAYKTFPPTEPKKDKKKKTDEAQPLPQRQKKQPT
              VTLLPAADMDDFSRQLQNSMSGASADSTQA"

ORIGIN

Query Match        100.0%;   Score 21;  DB 14;  Length 2304;
Best Local Similarity   100.0%;   Pred. No. 0.63;
Matches  21;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 ATGACCACACAAGGCAGATGG 21
          |||||||||||||||||||||
Db   2033 ATGACCACACAAGGCAGATGG 2053


RESULT 5
AY290752
LOCUS       AY290752         2810 bp    RNA       linear   VRL 09-FEB-2004
DEFINITION  SARS coronavirus ZJ01 isolate ZJ01b uncharacterized protein 6,
            uncharacterized protein 7a, uncharacterized protein 7b,
            uncharacterized protein 8a, uncharacterized protein 8b,
            nucleocapsid protein, uncharacterized protein 9b, and
            uncharacterized protein 9c genes, complete cds.
ACCESSION   AY290752
VERSION     AY290752.2  GI:38735509
KEYWORDS    .
SOURCE      SARS coronavirus ZJ01
  ORGANISM  SARS coronavirus ZJ01
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1 (bases 1 to 2810)
  AUTHORS   Li,L., Wang,Z., Lu,Y., Bao,Q., Chen,S., Wu,N., Cheng,S., Weng,J.,
            Zhang,Y., Yan,J., Mei,L., Wang,X., Zhu,H., Yu,Y., Zhang,M., Li,M.,
            Yao,J., Lu,Y., Yao,P., Bo,X., Wo,J., Wang,S. and Hu,S.
  TITLE     Severe acute respiratory syndrome-associated coronavirus genotype
            and its characterization
  JOURNAL   Chin. Med. J. 116 (9), 1288-1292 (2003)
   MEDLINE   22889812
   PUBMED    14527350
REFERENCE   2 (bases 1 to 2810)
  AUTHORS   Wang,Z.G., Li,L.J., Luo,Y., Zhang,J.Y., Wang,M.Y., Cheng,S.Y.,
            Zhang,Y.J., Wang,X.M., Lu,Y.Y., Wu,N.P., Mei,L.L. and Wang,Z.X.
  TITLE     Molecular biological analysis of genotyping and phylogeny of severe
            acute respiratory syndrome associated coronavirus
  JOURNAL   Chin. Med. J. 117 (1), 42-48 (2004)
   PUBMED    14733771
REFERENCE   3 (bases 1 to 2810)
  AUTHORS   Wang,Z., Cheng,S., Zhang,Y., Lu,Y. and Mei.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-MAY-2003) Department of Microbiology, Zhejiang Center
```

```
            KELSPRWYFYYLGTGPEASLPYGANKEGIVWVATEGALNTPKDHIGTRNPNNAATVL
            QLPQGTTLPKGFYAEGSRRGGSQASSRSSSRGNSRNSTPGSSRGNSPARMASGGGET
            ALALLLDRLNQLESKVSGKGQQQOGQTVTKKSAARASKKPRQKRTATKQYNVTQAFG
            RRGPEQTQGNFGQDQLIRQGTDYKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTY
            HGAIKLDDKDPQFKDNVILLNKHIDAYKTFPTPEPKKDKKKKTDEAQPLPQRQKKQPT
CDS                 VTLLPAADMDDFSRQLQNSMSGGASADSTQA"
            1218..1514
                     /codon_start=1
                     /product="uncharacterized protein 9b"
                     /protein_id="AAP69655.1"
                     /db_xref="GI:31747858"
                     /translation="MDPNQTNVVPPALHLVDPQIQLTITRMEDAMGQGQNSADPKVYP
                     IILRLGSQLSLSMARRNLDSLEARAFQSTPIVVQMTKLATTEELPDEFVVTAK"
CDS         1671..1883
                     /codon_start=1
                     /product="uncharacterized protein 9c"
                     /protein_id="AAP69660.1"
                     /db_xref="GI:31747859"
                     /translation="MLPPCYNFLKEQHCQKASTQREBAEAAVKPLLAPHHVVAVIQEIQ
                     LLAAVGEILLLEWLAEVVKLPSRYCC"
ORIGIN

Query Match             100.0%;  Score 21;  DB 14;  Length 2810;
Best Local Similarity   100.0%;  Pred. No. 0.63;
Matches    21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGACCACCAAGGCAGATGG 21
              ||||||||||||||||||||
Db       2486 ATGACCACCAAGGCAGATGG 2506

RESULT 6
AY304489
LOCUS       AY304489              8439 bp    RNA     linear   VRL 05-NOV-2003
DEFINITION  SARS coronavirus SZ1, partial genome.
ACCESSION   AY304489
VERSION     AY304489.1  GI:34482140
KEYWORDS    .
SOURCE      SARS coronavirus SZ1
  ORGANISM  SARS coronavirus SZ1
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 8439)
  AUTHORS   Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
            Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,
            Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and
            Poon,L.L.M.
  TITLE     Isolation and characterization of viruses related to the SARS
            coronavirus from animals in southern China
  JOURNAL   Science 302 (5643), 276-278 (2003)
  MEDLINE   22913660
   PUBMED   12958366
REFERENCE   2  (bases 1 to 8439)
  AUTHORS   Guan,Y. and Zheng,B.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAY-2003) Microbiology, The University of Hong Kong,
            University Pathology Building, Queen Mary Hospital, Pokfulam Road,
            Hong Kong, China
FEATURES             Location/Qualifiers
     source          1..8439
                     /organism="SARS coronavirus SZ1"
                     /mol_type="genomic RNA"
                     /isolate="SZ1"
                     /db_xref="taxon:231516"
                     /country="Hong Kong"
ORIGIN

Query Match             100.0%;  Score 21;  DB 14;  Length 8439;
Best Local Similarity   100.0%;  Pred. No. 0.68;
Matches    21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGACCACCAAGGCAGATGG 21
              ||||||||||||||||||||
```

```
Db        8268 ATGACCACCAAGGCAGATGG 8288

RESULT 7
AY304487
LOCUS       AY304487              8581 bp    RNA     linear   VRL 05-NOV-2003
DEFINITION  SARS coronavirus SZ13, partial genome.
ACCESSION   AY304487
VERSION     AY304487.1  GI:34482138
KEYWORDS    .
SOURCE      SARS coronavirus SZ13
  ORGANISM  SARS coronavirus SZ13
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 8581)
  AUTHORS   Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
            Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,
            Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and
            Poon,L.L.M.
  TITLE     Isolation and characterization of viruses related to the SARS
            coronavirus from animals in southern China
  JOURNAL   Science 302 (5643), 276-278 (2003)
  MEDLINE   22913660
   PUBMED   12958366
REFERENCE   2  (bases 1 to 8581)
  AUTHORS   Guan,Y. and Zheng,B.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-MAY-2003) Microbiology, The University of Hong Kong,
            University Pathology Building, Queen Mary Hospital, Pokfulam Road,
            Hong Kong, China
FEATURES             Location/Qualifiers
     source          1..8581
                     /organism="SARS coronavirus SZ13"
                     /mol_type="genomic RNA"
                     /isolate="SZ13"
                     /db_xref="taxon:231514"
                     /country="Hong Kong"

ORIGIN

Query Match             100.0%;  Score 21;  DB 14;  Length 8581;
Best Local Similarity   100.0%;  Pred. No. 0.68;
Matches    21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGACCACCAAGGCAGATGG 21
              ||||||||||||||||||||
Db       8262 ATGACCACCAAGGCAGATGG 8282

RESULT 8
AY304491
LOCUS       AY304491             11006 bp    RNA     linear   VRL 05-NOV-2003
DEFINITION  SARS coronavirus GZ60, partial genome.
ACCESSION   AY304491
VERSION     AY304491.1  GI:34482142
KEYWORDS    .
SOURCE      SARS coronavirus GZ60
  ORGANISM  SARS coronavirus GZ60
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 11006)
  AUTHORS   Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
            Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,
            Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and
            Poon,L.L.M.
  TITLE     Isolation and characterization of viruses related to the SARS
            coronavirus from animals in southern China
  JOURNAL   Science 302 (5643), 276-278 (2003)
  MEDLINE   22913660
   PUBMED   12958366
REFERENCE   2  (bases 1 to 11006)
  AUTHORS   Guan,Y. and Zheng,B.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAY-2003) Microbiology, The University of Hong Kong,
```

```
             University Pathology Building, Queen Mary Hospital, Pokfulam Road,
             Hong Kong, China
FEATURES             Location/Qualifiers
     source          1..11006
                     /organism="SARS coronavirus GZ60"
                     /mol_type="genomic RNA"
                     /isolate="GZ60"
                     /db_xref="taxon:231518"
                     /country="Hong Kong"
ORIGIN

Query Match             100.0%;  Score 21;  DB 14;  Length 11006;
Best Local Similarity   100.0%;  Pred. No. 0.69;
Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGACCACCAAGGCAGATGG 21
              ||||||||||||||||||||||
Db      10669 ATGACCACCAAGGCAGATGG 10689


RESULT 9
AY304493
LOCUS       AY304493               11010 bp    RNA     linear   VRL 05-NOV-2003
DEFINITION  SARS coronavirus HKU-65806, partial genome.
ACCESSION   AY304493
VERSION     AY304493.1  GI:34482144
KEYWORDS    .
SOURCE      SARS coronavirus HKU-65806
  ORGANISM  SARS coronavirus HKU-65806
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1 (bases 1 to 11010)
  AUTHORS   Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
            Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,
            Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and
            Poon,L.L.M.
  TITLE     Isolation and characterization of viruses related to the SARS
            coronavirus from animals in southern China
  JOURNAL   Science 302 (5643), 276-278 (2003)
  MEDLINE   22913660
   PUBMED   12958366
REFERENCE   2 (bases 1 to 11010)
  AUTHORS   Guan,Y. and Zheng,B.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAY-2003) Microbiology, The University of Hong Kong,
            University Pathology Building, Queen Mary Hospital, Pokfulam Road,
            Hong Kong, China
FEATURES             Location/Qualifiers
     source          1..11010
                     /organism="SARS coronavirus HKU-65806"
                     /mol_type="genomic RNA"
                     /isolate="HKU-65806"
                     /db_xref="taxon:231520"
                     /country="Hong Kong"
ORIGIN

Query Match             100.0%;  Score 21;  DB 14;  Length 11010;
Best Local Similarity   100.0%;  Pred. No. 0.69;
Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGACCACCAAGGCAGATGG 21
              ||||||||||||||||||||||
Db      10673 ATGACCACCAAGGCAGATGG 10693


RESULT 10
AY304494
LOCUS       AY304494               11010 bp    RNA     linear   VRL 05-NOV-2003
DEFINITION  SARS coronavirus HKU-66078, partial genome.
ACCESSION   AY304494
VERSION     AY304494.1  GI:34482145
KEYWORDS    .
SOURCE      SARS coronavirus HKU-66078
  ORGANISM  SARS coronavirus HKU-66078
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1 (bases 1 to 11010)
  AUTHORS   Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
            Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,
            Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and
            Poon,L.L.M.
  TITLE     Isolation and characterization of viruses related to the SARS
            coronavirus from animals in southern China
  JOURNAL   Science 302 (5643), 276-278 (2003)
  MEDLINE   22913660
   PUBMED   12958366
REFERENCE   2 (bases 1 to 11010)
  AUTHORS   Guan,Y. and Zheng,B.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAY-2003) Microbiology, The University of Hong Kong,
            University Pathology Building, Queen Mary Hospital, Pokfulam Road,
            Hong Kong, China
FEATURES             Location/Qualifiers
     source          1..11010
                     /organism="SARS coronavirus HKU-66078"
                     /mol_type="genomic RNA"
                     /isolate="HKU-66078"
                     /db_xref="taxon:231521"
                     /country="Hong Kong"
ORIGIN

Query Match             100.0%;  Score 21;  DB 14;  Length 11010;
Best Local Similarity   100.0%;  Pred. No. 0.69;
Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGACCACCAAGGCAGATGG 21
              ||||||||||||||||||||||
Db      10673 ATGACCACCAAGGCAGATGG 10693


RESULT 11
AY304490
LOCUS       AY304490               13471 bp    RNA     linear   VRL 05-NOV-2003
DEFINITION  SARS coronavirus GZ43, partial genome.
ACCESSION   AY304490
VERSION     AY304490.1  GI:34482141
KEYWORDS    .
SOURCE      SARS coronavirus GZ43
  ORGANISM  SARS coronavirus GZ43
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1 (bases 1 to 13471)
  AUTHORS   Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
            Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,
            Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and
            Poon,L.L.M.
  TITLE     Isolation and characterization of viruses related to the SARS
            coronavirus from animals in southern China
  JOURNAL   Science 302 (5643), 276-278 (2003)
  MEDLINE   22913660
   PUBMED   12958366
REFERENCE   2 (bases 1 to 13471)
  AUTHORS   Guan,Y. and Zheng,B.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAY-2003) Microbiology, The University of Hong Kong,
            University Pathology Building, Queen Mary Hospital, Pokfulam Road,
            Hong Kong, China
FEATURES             Location/Qualifiers
     source          1..13471
                     /organism="SARS coronavirus GZ43"
                     /mol_type="genomic RNA"
                     /isolate="GZ43"
                     /db_xref="taxon:231517"
                     /country="Hong Kong"
ORIGIN
```

```
Query Match            100.0%;   Score 21;   DB 14;   Length 13471;
Best Local Similarity  100.0%;   Pred. No. 0.7;
Matches   21;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy        1 ATGACCACAAGGCAGATGG 21
            ||||||||||||||||||||
Db    13134 ATGACCACAAGGCAGATGG 13154

RESULT 12
AY304492
LOCUS       AY304492               13471 bp    RNA     linear   VRL 05-NOV-2003
DEFINITION  SARS coronavirus HKU-36871, partial genome.
ACCESSION   AY304492
VERSION     AY304492.1  GI:34482143
KEYWORDS    .
SOURCE      SARS coronavirus HKU-36871
  ORGANISM  SARS coronavirus HKU-36871
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 13471)
  AUTHORS   Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,
            Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,
            Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S.M. and
            Poon,L.L.M.
  TITLE     Isolation and characterization of viruses related to the SARS
            coronavirus from animals in southern China
  JOURNAL   Science 302 (5643), 276-278 (2003)
  MEDLINE   22913660
   PUBMED   12958366
REFERENCE   2  (bases 1 to 13471)
  AUTHORS   Guan,Y. and Zheng,B.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAY-2003) Microbiology, The University of Hong Kong,
            University Pathology Building, Queen Mary Hospital, Pokfulam Road,
            Hong Kong, China
FEATURES             Location/Qualifiers
     source          1..13471
                     /organism="SARS coronavirus HKU-36871"
                     /mol_type="genomic RNA"
                     /isolate="HKU-36871"
                     /db_xref="taxon:231519"
                     /country="Hong Kong"
ORIGIN

Query Match            100.0%;   Score 21;   DB 14;   Length 13471;
Best Local Similarity  100.0%;   Pred. No. 0.7;
Matches   21;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy        1 ATGACCACAAGGCAGATGG 21
            ||||||||||||||||||||
Db    13134 ATGACCACAAGGCAGATGG 13154

RESULT 13
AY463060
LOCUS       AY463060               29013 bp    RNA     linear   VRL 05-JAN-2004
DEFINITION  SARS coronavirus ShanghaiQXC2, complete genome.
ACCESSION   AY463060
VERSION     AY463060.1  GI:40457448
KEYWORDS    .
SOURCE      SARS coronavirus ShanghaiQXC2
  ORGANISM  SARS coronavirus ShanghaiQXC2
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 29013)
  AUTHORS   Yuan,Z., Zhang,X., Hu,Y., Lan,S., Zhou,Z., Wang,H. and Wen,Y.
  TITLE     Analysis of SARS coronavirus genome in Shanghai isolates
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 29013)
  AUTHORS   Yuan,Z., Zhang,X., Hu,Y., Lan,S., Zhou,Z., Wang,H. and Wen,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-NOV-2003) Key Lab of Medical Molecular Virology,
```

Shanghai Medical College, Fudan University, 138 Yi Xue Yuan Road,
Shanghai 200032, P.R. China

```
FEATURES             Location/Qualifiers
     source          1..29013
                     /organism="SARS coronavirus ShanghaiQXC2"
                     /virion
                     /mol_type="genomic RNA"
                     /strain="ShanghaiQXC2"
                     /db_xref="taxon:258508"
     CDS             join(186..12734,12734..20827)
                     /note="ribosomal slippage"
                     /codon_start=1
                     /product="orf1ab"
                     /protein_id="AAR86774.1"
                     /db_xref="GI:40457449"
                     /translation="MESLVLGVNEKTHVQLSLPVLQVRDVLVRGFGDSVEEALSEARE
                     HLKNGTCGLVELEKGVLPQLEQPYVFIKRSDALSTNHGHKVVELVAEMDGIOYGRSGI
                     TLGVLVPHVGETPIAYRNVLLRKNGNKGAGHSYGIDLKSYDLGDELGTDPIEDYEQN
                     WNTKHGSGALRELTRELNGGAVTRYVDNNFCGPDGYPLDCIKDFLARAGKSMCTLSEQ
                     LDYIESKRGVVCCRDHEIAWFTERSDKSYEHQTPEIKSAKKFDTFKGECPKFVFP
                     LNSKVKVIQPRVEKKKTEGFMGRIRSVYPVASPQECNNMHLSTLMKCNHCDEVSWQTC
                     DFLKATCEHCGTENLVLEGPTTCGYLPTNAVVKMPCACQDPEIGPEHSVADYHHSN
                     IETRLRKGGRTRCFGGCVFAVVGCYNKRAYWVPRASADIGSGHTGITGDNVETLNEDL
                     LEILSRERVNINIVGDFHLNEEVAILIASFSASTSAFIDTIKSLDYKSFXTIVESCGN
                     YKVTKGKPVKGAWNIGQQRSVLTPLCGFPSQAGAVIRSIFARTLDAANHSIPDLQRAA
                     VTILDGISEQSLRLVDAMVYTSDLLTNSVIIMAYVTGGLVQOTSQWLSNLLGTTVEKL
                     RPIFEWIEAKLSAGVEFLKDAWEILKFLITGVFDIVKGQIQVASDNIKDCVKCFIDVV
                     NKALEMCIDQVVIAGAKLRSLNLGEVFIAQSKGLYRQCIRGKEQLQLLMPLKAPKEVT
                     FLEGDSHDTVLTSEEVVLKNGELEALETPVDSFTNGAIVGTPVCVNGLMLLEIKDKEQ
                     YCALSPGLLATNNVFRLKGGAPIKGVTFGEDTVWEVGQYKNVRITFELDERVDKVLNE
                     KCSVYTVESGTTVTEFACVVAAVKTLQPVSDLLTNMGIDLDEWSVATFYLFDDAGE
                     ENFSSRMYCSFYPPDEEEEDDAECEEEIDETCEHEYGTEDDYQGLPLEFGASAETVR
                     VEEEEEDWLDDTTEQSEIEPEPEPTPEEPVNQFTGYLKLTDNVAIKCVDTVKEAQSA
                     NPMVIVNAANIHLKHGGVAGALNKATNGAMQKESDDYIKLNGPLTVGGSCLLSGHNL
                     AKKCLHVVGPNLNAGEDIQLLKAAYENFNSQDILLAPLLSAGIFGAKPLQSLQVCVQT
                     VRTQVVIAVNDKALYEQVVMDYLDNLKPRVEAPKQEEPPNTEDSKTEEKSVVQKPVNH
                     KPKIKACIDEVTTLEETKFLTNKLLLLFADINGKLYHDSQNMLRGEDMSFLEEDAPYM
                     VGDVITSGDITCVIPSKKAGGTTEMLSRALKKVPVDEYITTYPGQGCAGYTLEEAKT
                     ALKKCKSAFYVLPSEAPNAKEEILGTVSWNLREMLAHAEEARKLMPICMDVRAIMATI
                     QRKYKVGVKIQEGIVDYGVRFFYTSKEPVASIITKLNSLNEPLVTMPIGYVTHGFNLE
                     EAARCMRSLKAPAVVSVSSPAVTTYNGYLTSSKTSEEHFVETVSLAGSYRDWSYSG
                     QRTELGVEFLKRGDKIVYHTLESPVEFHLDGEVSLDKLSLLSLREVKTIKVFTTVD
                     NTNLHTQLVDMSMTYGQQFGPTYLDGADVTKIKPHVNHEGKTFFVLPSDDTLRSEAFE
                     YYHTLDESFLGRYMSALNHTKKWKFPQVGGLTSIKWADNNCYLSSVLLALQOLEVKFN
                     APALQAYYRARAGDAANFCALILAYSNKTVGELGDVRETMHLLQHANLESAKRVLN
                     VCKHCGQKTTLTGVEAVMYMGTLSYDNLKTGVSIPCVCGRDATQYIVQQSSFVMM
                     SAPPAGYKLQQGTFLCANEYTGNYQCGHYTHITAKETLYRIDGAHLTKMSEYKGPVTD
                     VPYKESYTTIKPVSYKLDGVTYTEIEPKLDGYYKKDNAYVENTSITIKKPNELSLA
                     LGLKTIATHGIAINSVPWSKILAYVKPFLGQAAITTSNCAKRLAQRVFNNIMPYVFT
                     LLFQLCTFTKSTNSRIRASLPTTIAKNSVKSVAKLCLDAGINYVKSPKFSKLFTIAMW
                     LLLSICLGSLICVTAAFGVLLSNFGAPSYCNGVRELYLNSSNVTTMDFCEGSFPCSI
                     CLSGLDSLDSYPALETIQVTISSYKLDLTILGLAAEWVLAYMLFTKFFYLLGLSAIMQ
                     AFFGYFASHFISNSWLMWFIISIVQMAPVSAMVRMYIFFASFYYIWKSYVHIMDGCTS
                     STCMMCYKCNHATRVECTTIVNGMKRSFYVANGGRGFCKTHNWNCLNCDTFCTGSTF
                     ISDEVARDLSLQFKRPINPTDQSSYIVDSVAVNGALHPYFDKAGQKTYERHPLSHFV
                     NLDNLRANNTKGSLPINVIVFDGKSKCDESASKSASVYYSQLMCQPILLLDQALVSDV
                     GDSTEVSVKMFDAYVDTFSAIFSVPMEKLKALVATAHSEAKAVLTVSSAALSTFVSAAR
                     QGVVDTDVDTKDVIECLKLSHHSDLEVTGDSCNFMJTYNKVENMTPRDLGACIDCNA
                     RHINAQVAKSHNVSLIWNVKDYMSLSEQLRKQIRSAAKKNNIPFRLTCATTRQVVNVI
                     TTKISLKGGKIVSTCFKLMLKATLLCVLAALVCYIVMPVHTLSIHDGYTNEIIGYKAI
                     QDGVTRDIISTDDCFANKHAGPDAWFSQRGGSYKNDKSCPVVAAIITREIGFIVPGLP
                     GTVLRAINGDFLHFLPRVFSAVGNICYTPSKLIEYSDFATSACVLAAECTIFKDAMGK
                     PVPYCYDTNLLEGSISYSELRPDTRVVLMDGSIIQFPNTYLEGSVRVVTTFDAEYCRH
                     GTCERSEVGICLJSTSGRWVLNNEHYRALSGVFCGVDAMNLIANIFTPLVQPVGALDAS
                     ASVVAGGIIAILVTCAAYYFMKFRRAFGEYNHVVAANALLFLMSFTILCLVPAYSFLP
                     GVYSVFYLYLTFYFTNDVSFLAHLQWFAMFSPIVPFWITAIVYFCISLKCHWFFNNY
                     LRKRVMFNGVTFSTFEEAALCTFLLNKEMYLKLRSETLLPLTQYNRYLALYNKYKYFS
                     GVLDTTSYREAACCHLAKALNDFSNSGADVLYQPPQTSITSAVLQSGFRKMAFFSGKV
                     EGCMVDTVCGTTLGWLLDDTVYCPRHVICTAEDMLNPNYEDLLIRKSNHSFLVQAG
                     NVQLRVIGHSMQNCLLRLKVDTISNPKTPKYKFVRIQPGQTFSVLACYNGSPSGVYQCA
                     MRPNHTIKGSFLNGSCGSVGFNIDYDCVSFCYMHHMELPTGVHAGTDLEGKFYGPFVD
                     RQTAQAAGTDTTITLNVLAWLYAAVINGDRWFLNRFTTLNDFNLVAMKYNYEPLTQD
                     HVDILGPLSAQTGIAVLDMCAALKELLQNGMNGRTILGSTILEDEFTPFDVVRQCSGV
```

```
DVRQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRLRHGKLRP
FERDISNVFPSPDGKPCTPALNCYWPLNDYGFYTTGIGYQPYRVVLSFELLNAPA
TVCGPKLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRQFPQFGQFGRRDVSDFTDSVRDPK
TSEILDISPCAFGGVSVITPGTNASSEVAVLYQDVNCTNVSAAIHADQLTPAWRIYST
GNNVFQTQAGCLIGAEHVDTSYECDIPIGAGICASYHTVSLLRSTSQKSIVAYTMSLG
ADSSIAYSNNTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGS
FCTCLNRALSGIAAEQDRNTREVFAQVKQMYKTPTLKYFGGFNFSQILPDPLKPTKRS
FIEDLLFNKVTLADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLITDMIAAYT
AALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAIS
QIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAE
VQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYH
LMSFPQAAPHGVVFLHVTYVPSQERNFTTAPAICHEGKAYFPREGVFVFNGTSWFITQ
RNFFSPQIITTDNTFVSGNCDVVIGIINNTVYDPLQPELDSFKGELDKYFKNHTSPDV
DLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIA
GLIAIVMVTILLCCMTSCCSCLKGACSCGSCCKFDEDDSEPVLKGVKLHYT"
24610...25434
/codon_start=1
/product="unknown"
/protein_id="AAR86776.1"
/db_xref="GI:40457451"
/translation="MDLFMRFFTLGSITAQPVKIDNASPASTAHATATIPLQASLPFG
WLVIGVAFLAVFQSATKIIALNKRWQLALYKGFOICNLLLFVTIYSHLLVAAGME
AQFLYLYALLYFLQCINACRIIMRCWLCWKCKSKNPLLYDANYFVCWHTHNYDYCIPY
NSVTDTIVVTEGGDISTPRALKEDYIGGYSEDRHSGVDCYVVHGYFTEVYLQLESTQ
ITNNTGIENATFFIFNKLVKDPPNVQIHTIDGSSGVANPAMDPIYDEPTTTTSVPL"
25031...25495
/codon_start=1
/product="unknown"
/protein_id="AAR86777.1"
/db_xref="GI:40457452"
/translation="MMPTTLFAGTHITMTVYHITVSQIQLSLLKVTAFQHQNSKTT
KLIVTLRIGTQVLKTMSLYMAISPKFTTSLSLHKLLQTLVKLMLHSSSLTSLLKTHRM
CKYTQSTALQELLIQQWIQFMMSRRRLLACLCKHKKVSTNLCTHSFRKKQVR"
25459...25689
/codon_start=1
/product="E"
/protein_id="AAR86778.1"
```

```
Query Match      100.0%;     Score 21;     DB 14;     Length 29013;
Best Local Similarity    100.0%;     Pred. No. 0.73;
Matches     21;   Conservative    0;    Mismatches    0;    Indels    0;    Gaps    0;

Qy      1  ATGACCACAAGGCAGATGG  21
            |||||||||||||||||||||
Db   28740  ATGACCACAAGGCAGATGG  28760
```

```
RESULT 14
AY394999
LOCUS       AY394999              29350 bp    RNA     linear   VRL 29-JAN-2004
DEFINITION  SARS coronavirus LC2, complete genome.
ACCESSION   AY394999
VERSION     AY394999.1  GI:37624342
KEYWORDS    .
SOURCE      SARS coronavirus LC2
  ORGANISM  SARS coronavirus LC2
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 29350)
  AUTHORS   The SARS epidemiology consortium of Guangdong
  CONSRTM   From independent foci of epidemic outbreak to large genomic
  TITLE     alteration in late phase viruses: evolution of the SARS-coronavirus
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 29350)
  AUTHORS   The SARS epidemiology consortium of Guangdong
  CONSRTM   Direct Submission
  TITLE     Submitted (19-SEP-2003) Guangdong, China
  JOURNAL   Location/Qualifiers
FEATURES    1..29350
            /organism="SARS coronavirus LC2"
            /mol_type="genomic RNA"
     source
```

```
TFGQKFKKIVKGTHHWMLLTFLTSLLILVQSTQWSLFFVYENAFLPFTLGIMAIAAC
AMLLVKHKHAFLCLFLLPSLATVAYFNMVYMPASWVMRIMTWLELADTSLSGVRLKDC
VMTASALVLLILMARTVVDDAARRWTIMNVTLVVKYYGNALDQAISMALVISA
TSNYSGVVTTIMFLARA1VFVCVEYYPLLFITGNTLQCIMLVVCFLGYCCCYFGLFC
LLNRYFRLTLGVVDYLVSTQEFRYMNSQCLLPKKSSIDAFKLNIKLLGIGGKFCIKVA
TVQVSKMSDDVKCTSVVLLSLQQLRVESSSKLWACQVQLHNDILLAKDTTEAFEKMVSL
LSVLLSMQGAVDINRLCEEVLDNRAATLQAIASEPFSSLPSYAAYATAQEAYEQAVANGD
SEVVLKKLKKSLNVAKSFRDRAAMQRKLEKMADQAWTQMYKQARSEDKRAKVTSAMQ
TMLFTMRLRDNDALNNIINNARDYCLPNIIPLTTAAKLMVVVPDYGTKYCTCDGNT
FTYASALWEIOQVVDADSKIVQLSEIMDNSPNLAWPLIVTALRANSAAKLQNNELSP
VALRQMSCAAGTTQTACTDDNALAYYNNSKGGRFVLALLSDHQDLKWARFFKSDGTGT
IYTELEPPCRFVVTDTVQGPKKMKYLYFIKGLNNLNRGMVGLSLAATVRLQAGNATVPA
NSTVLSFCAFAVDPAKAYKDYLASGGQPITNCVKMLCTHTGTSQAITVTPEARMDQGS
FGGASCCLYCRCHIDHPNPKGFCDLKGKYVQIPTTCANDPVGFTLRNTVCTVCGMWKG
YGCSCDQLREPLMQSADASTFFFKRVCGVSAARLTPCGTGTSTDVVYRAFDIYNEKVAG
FAKFLKTNCCRFQEKDEEGNLLDSYFVVKRHTMSNYQHEETIYNLIKDCPAVAVHDFF
KFRVDGDMVPHISRQRLIKTYMADLVYALRHFDEGNCDTLKEILVTYNCCDDYFNKK
DWYDFVENPDILRYVANLGERVRQSLLKTVQFCDAMRDAGIVGVLTLNDQDLNGNWYD
FGDFVQVAPGCGVPIVDSYYSLLMPILTLTRALLAAESHMDADLAKPLIKWDLLKYDFT
EERLCLFDRYFKYWDQTTHPNCINCLDDRCILHCANFNVLFSTVFPPTSFGPLVRKIF
VDGVFVVSTGYHFRELGVVHNQDVNLHSRLSPKELLVYAADPAMHAASGNLLVHA
TTCFSVAALTNNVAFQTVKPGNFNKDFYDFAVSKGFFKEGSSVELKHFFAQDGNAAI
SDYDYRYNLPTMCDIRQLLFVEVVDKYFDCYDGGCINANQVIVNNLDKSAGFPFNK
WGKARLYDSMSYEDQDALFAYTKRNVIPTITOWMLKYAISANQFRMPTVAGVSICSTM
TNRQFHQKLLKSIAATRGATVVIGTSKFYGGWHNMLKTVYSDVETPHLMGWDYPKCDR
AMPNMLRIMASLVLARKHNTCCNLSHRFYRLANECAQVLSEMVMCGGSLYVKPGGTSS
GDATTAYANSVFNICQAVTANVNALLSTDGNKIADKVYRNLQHRLYECLYRNRDVDHE
FVDEFYAYLRKHFSMMILSDDAVVCYNSNYAAOGLVASIKNFKAVLYYQNNVFMSEAK
CWTETDLYKGPHEFCSQHTMLVKQGDDYYLPYPDPSRILGAGCFVDDIVKTDGTLMI
ERFVSLAIDAYPLTKHPNQEYADVFHLYLQYIRKLHDELTGHMLDMYSVMLTNDNTSR
YWEPEFYEAMYTPHTVLQAVGACVLCNSQTSLRCGACIRRPFLCCKCCYDHVISTSHK
LVLSVNPYVCNAPGCDVTDVTQLYLGGMSYYCKSHKPISFFLCANGQVFGLYKNTCV
GSDNVTDFNAIATCDWTNAGDYIIANTCTERLKLFAAETLKATEETFKLSYGIATVRE
VLSDRELHLSWEVGKPRPPLNRNYVFTGYRVTNSKVQIGEYTFEKGDYGDAVVYRGT
GMQKYSTLQGPPGTGKSHFAIGLALYYPSARIVTACSHAVDCALCELAKYLPIDKC
SRIIPARARVECFDKFKVNSTLEQYVFCTVNALPETTADIVVFDEISMATNYDLSVVN
ARLRAKHYVYIGDPAQLPAPRTLLTKGTLEPEYFNSVCRLMKTIGPDMFLGTCRRCPA
EIVDTVSALIVQNKVLKAHKEKAQYKVTHDVSSAINRPQIGVVREFLTRNP
AWRKAVFISPYNSQNAVASKILGLPTQVTVDSSQGSEYDVIFTQTTETAHSCNVNRFN
VAITRAKIGILCIMSDRDLYDKLQFTSLEIPRRNVATLQAENVTGLFKDCSKIITGLH
PTQAPTHLSVDIKFKTEGLCVDIPGIPKDMTYRRLISMGFKMNYQVNGYPNMFITRE
EAIRHVRAWIGFDVEGCHATRDAVGTNLPLQLGFSTGVNLVAVPTGYVDTENNTEFTR
NAKPPPGDQFKHLIPLMYKGLPWNVRIKIVQMLSDTLKGLSDRVVFVLWAHGFELT
SMKYFVKIGPERTCCLCDKRATCFSTSSDTYACWNHSVGFDYVYNPFMIDVQQWGFTG
NLQSNHDQHCQVHGNAHVASCDAIMTRCLAVHECFVKRVDWSVEYPIIGDELRVNSAC
RKVQHMVVKSALLADKFPVLHDIGNPKAIKCVPQAVEWVKFYDAQPCSDKAYKIEELF
YSYATHHDKFTDGVCLFWNCNVDRYPANAIVCRFDTRVLSNLNLPGCDGGSLYVNKHA
FHTPAPDKSAFTNLKQLPFFYYSDSPCESHGKQVVSDIDYVPLKSATCITRCNLGGAV
CRHHANEYRQYLDAYNMMISAGFSLWIYKQFDTYNLWNTFTRIQSLENVAYNVVNKGH
FDGHAGEAPVSIINNAVTKVDGLDVEIFESKTTLPVNVAFELWAKRNIKPVPEIKIL
NNLGVDIAANTVIWDYKREAPAHVSTIGVCTMTDIAKKPTESACSSLTVLFDGRVEGQ
VDLFRNARNGVLITEGSVKGLTFSKGPAQASVNGVTLIGESVKTQFNVFKKVDGIIQQ
LPETYFTQSRDLEDFKPRSQMETDFLELAMDEFIQRYKLEGYAFEHIVYGDFSHGQLG
GLHLMIGLAKRSQDSPLKLEDFIPMDSTVKNYITDAQTGSSKCVRSVIDLLLDDFVE
IIKSQDLSVISKVKVTIDYAEISFMLWCKDGHVETFYPKLQASQAWQPGVAMPNLYK
MQRMLLEKCDLQNYGENAVIPKGIMMNVAKYTQLCQYLNTLTLAVPYNMRVIHFGAGS
DKGVAPGTAVLRQWLPTGTLLVDSDLNDFVSDADSTLIGDCATVHTANKWDLIISDMY
DPRTKHVTKENDSKEGFFTYLCGFIKQKLALGGSIAVKITEHSWNADLYKLMGHFSWW
TAFVTNVNASSEAFLIGANYLGKPKEQIDGYTMHANYIFWRNTNPIQLSSYSLFDMS
KFPLKLRGTAVMSLKENQINDMIYSLLEKGRLLIRENNRVVVSSDILVNN"
20834...24601
/codon_start=1
/product="spike"
/protein_id="AAR86775.1"
/db_xref="GI:40457450"
/translation="MFIFLLFLTLTSGGSDLDRCTTFDDVQAPNYTQHTSSMRGVYYPD
EIPRSDTYLYTQDLLPFYSNVTGFHTINHTGNPVIPFKDGIYFAATEKSNVVRGWV
FGSTMNNKSQSVIINNSTNVVIRACNFELCDNPFFAVSKPMGTQTHTMIFDNAFNCT
FEYISDAFSLDVSEKSGNFKHLRFVFKNKDGFLYVYKGYPIDVVRDLPSGFNTLKP
IFKLPLGINITNFRAILTAFSPAQDIWGTSAAAYFVGYLKPTTFMLKYDENGTITDAV
DCSQNPLAELKCSVKSFEIDKGIYQTSNFRVPSGDVVRFPNITNLCPFGEVFNATKF
PSVYAWERRKISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCFSNVYADSFVVKGD
```

ORIGIN

```
                    /isolate="LC2"
                    /db_xref="taxon:249084"

Query Match
Best Local Similarity   100.0%;    Score 21;  DB 14;   Length 29350;
Matches      21;  Conservative  0;  Mismatches  0;  Indels    0;   Gaps    0;

Qy          1 ATGACCACCAAGGCAGATGG 21
              |||||||||||||||||||||
Db      28997 ATGACCACCAAGGCAGATGG 29017
```

```
RESULT 15
AY395000
LOCUS       AY395000              29350 bp    RNA      linear   VRL 29-JAN-2004
DEFINITION  SARS coronavirus LC3, complete genome.
ACCESSION   AY395000
VERSION     AY395000.1  GI:37624343
KEYWORDS    .
SOURCE      SARS coronavirus LC3
  ORGANISM  SARS coronavirus LC3
            Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
            Coronaviridae; Coronavirus.
REFERENCE   1  (bases 1 to 29350)
  AUTHORS   .
  CONSRTM   The SARS epidemiology consortium of Guangdong
  TITLE     From independent foci of epidemic outbreak to large genomic
            alteration in late phase viruses: evolution of the SARS-coronavirus
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 29350)
  AUTHORS   .
  CONSRTM   The SARS epidemiology consortium of Guangdong
  TITLE     Direct Submission
  JOURNAL   Submitted (19-SEP-2003) Guangdong, China
FEATURES             Location/Qualifiers
     source          1..29350
                     /organism="SARS coronavirus LC3"
                     /mol_type="genomic RNA"
                     /isolate="LC3"
                     /db_xref="taxon:249085"
```

ORIGIN

```
Query Match
Best Local Similarity   100.0%;    Score 21;  DB 14;   Length 29350;
Matches      21;  Conservative  0;  Mismatches  0;  Indels    0;   Gaps    0;

Qy          1 ATGACCACCAAGGCAGATGG 21
              |||||||||||||||||||||
Db      28997 ATGACCACCAAGGCAGATGG 29017
```

Search completed: May 15, 2005, 12:53:28
Job time : 95.9067 secs